

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:31:32 ; Search time 5716 Seconds  
(without alignments)  
10491.575 Million cell updates/sec

Title: US-10-510-408-1  
Perfect score: 1055  
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 583141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sv.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055	100.0	1055	6	AX930045 Sequence
2	1055	100.0	1522	6	AX954657 Sequence
c 3	1047.4	99.3	110000	1	Continuation (28 o
c 4	1047.4	99.3	110000	1	Continuation (28 o
5	459	43.5	459	6	AX434675 Sequence
6	457	43.3	17787	1	Y15896 Bacillus su
7	457	43.3	200707	1	BSUB0015
8	420.6	39.9	2664	1	BSBOPCGEN
c 9	209.6	19.9	110000	1	Continuation (26 o
c 10	209.6	19.9	110000	1	Continuation (27 o
11	195.4	18.5	546	6	AX433201
c 12	166	15.7	301332	1	AE017012
c 13	164.4	15.6	288514	1	AE017278
c 14	161.2	15.3	110000	1	Continuation (43 o
c 15	161.2	15.3	110000	1	Continuation (43 o
c 16	161.2	15.3	290029	1	AE017038
c 17	159.6	15.1	110000	1	Continuation (43 o
c 18	159.6	15.1	110000	1	Continuation (43 o

19	137	13.0	595	6	AX437557	AX437557 Sequence
c 20	132.8	12.6	290507	1	AE017327	AE017327 Listeria
21	132.2	12.5	110000	1	AP006627	Continuation (17 o
22	130.8	12.4	110000	1	BA000004_13	Continuation (14 o
c 23	124.8	11.8	240050	1	AL591979	AL591979 Listeria
c 24	124.8	11.8	349980	6	AX641669	AX641669 Sequence
c 25	122	11.6	239050	1	AL596169	AL596169 Listeria
c 26	122	11.6	349980	6	AX413017	AX413017 Sequence
c 27	122	11.6	349980	6	AX417045	AX417045 Sequence
28	107.6	10.2	300478	1	AE017201	AE017201 Lactobaci
29	107.6	10.2	349980	6	AX926712	AX926712 Sequence
c 30	106.8	10.1	110000	1	BA000028_20	Continuation (21 o
c 31	103.4	9.8	110000	1	CP000029_12	Continuation (13 o
c 32	103.4	9.8	300029	1	AE016748	AE016748 Staphyloc
33	98.6	9.3	110000	1	AP008934_11	Continuation (12 o
c 34	98.4	9.3	297050	1	AL935258	AL935258 Lactobaci
35	98.2	9.3	110000	1	CP000033_03	Continuation (4 of
c 36	94.2	8.9	11654	1	AE015030	AE015030 Streptoco
37	91.4	8.7	175936	14	SPNEU1908	AL449930 Streptoco
38	89.8	8.5	11344	1	AE007333	AE007333 Streptoco
39	89.8	8.5	14273	6	CQ788949	CQ788949 Sequence
40	89.8	8.5	14273	6	AR218808	AR218808 Sequence
41	89.8	8.5	14273	6	BD003720	BD003720 Polynucle
42	89.8	8.5	349980	6	AX571760	AX571760 Sequence
43	89	8.4	591	6	AX566968	AX566968 Sequence
44	89	8.4	609	6	AX617338	AX617338 Sequence
45	89	8.4	7159	6	AR354184	AR354184 Sequence

ALIGNMENTS

RESULT 1  
AX930045  
LOCUS AX930045 1055 bp DNA linear PAT 22-DEC-2003  
DEFINITION Sequence 1 from Patent WO03087148.  
ACCESSION AX930045  
VERSION AX930045.1 GI:40311862  
KEYWORDS  
SOURCE Bacillus licheniformis  
ORGANISM Bacillus licheniformis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
REFERENCE 1  
AUTHORS Andersen,J.T., j Rgensen,S.T., Rasmussen,M.D., Olsen,P.B. and Clausen,I.G.  
TITLE Improved bacillus host cell  
JOURNAL Patent: WO 03087148-A 1 23-OCT-2003;  
Novozymes A/S (DK)  
FEATURES  
source  
Location/Qualifiers  
1..1055  
/organism="Bacillus licheniformis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1402"  
34..555  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF03678.1"  
/db\_xref="GI:40311863"  
/translation="MYSRKFKIGILLIGSLAALSFHLEALAKPAKVQIQLEKYYL  
DGDYIENKVEARTLDFKAYGWLIDQKGFILFRKQVDDISPLSKNTGYIGVT  
EDGVISYFPHGRPGILSEPIQSFQIDIKRLSRMADDLRKGIPYRTKKEFEHVEAVK  
SSGSQHVHEDMKT"

ORIGIN

Query Match 100.0%; Score 1055; DB 6; Length 1055;  
Best Local Similarity 100.0%; Pred. No. 6.5e-279;  
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCTTAAGCTAGGCAAAAAGAGCTGATTTGATGTCAGCCGCAAGCAAGTTCAAAATC 60  
DB 1 CCTTAAGCTAGGCAAAAAGAGCTGATTTGATGTCAGCCGCAAGCAAGTTCAAAATC 60

```
QY 61 GGTATTATGCTTATTGGAAGTCTGCTGCGCGCTCAGCTTTCACCTGGAGCGCTTGCGCC 120
DB 61 GGTATTATGCTTATTGGAAGTCTGCTGCGCGCTCAGCTTTCACCTGGAGCGCTTGCGCC 120
QY 121 GAAAGCGCGCTAAAGTTCAAAATCAGCTTGAAAGGTTTATCTGCGAGGAGCGTTGGA 180
DB 121 GAAAGCGCGCTAAAGTTCAAAATCAGCTTGAAAGGTTTATCTGCGAGGAGCGTTGGA 180
QY 181 ATTGAGAAATAAGTAGAGCGCGCTCGCACCTGGAAGACTTTTAAAGTGTCTTATAAAGGG 240
DB 181 ATTGAGAAATAAGTAGAGCGCGCTCGCACCTGGAAGACTTTTAAAGTGTCTTATAAAGGG 240
QY 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTCGACGACATT 300
DB 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTCGACGACATT 300
QY 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 360
DB 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 360
QY 361 TTTCACGCTCGCGGGCATCTTATCAGAAACCAATTCATCGTTTTCAGATTGATATA 420
DB 361 TTTCACGCTCGCGGGCATCTTATCAGAAACCAATTCATCGTTTTCAGATTGATATA 420
QY 421 AAGCGCTGGAAGCGGATCGCGATGATCTGCGCAAGGATACCATACCGCACGAA 480
DB 421 AAGCGCTGGAAGCGGATCGCGATGATCTGCGCAAGGATACCATACCGCACGAA 480
QY 481 AAGGAATTGAACATGTCATTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 540
DB 481 AAGGAATTGAACATGTCATTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 540
QY 541 GATATGAAGACATGACGCTGTTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 600
DB 541 GATATGAAGACATGACGCTGTTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 600
QY 601 AACATATGTTAACTTTTTCATTTAGCTTTGCTGCTTTGTTGTTTACATGAAGAGCAGTC 660
DB 601 AACATATGTTAACTTTTTCATTTAGCTTTGCTGCTTTGTTGTTTACATGAAGAGCAGTC 660
QY 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAGGAGGACGATTGATGATTCGCCCC 720
DB 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAGGAGGACGATTGATGATTCGCCCC 720
QY 721 AATATATGTCATTTGAAACCGCGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTA 780
DB 721 AATATATGTCATTTGAAACCGCGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTA 780
QY 781 TTTATAGAAAACAGCAAGAAACAACTATATACATACCATTTATGTAAGAGACACGCA 840
DB 781 TTTATAGAAAACAGCAAGAAACAACTATATACATACCATTTATGTAAGAGACACGCA 840
QY 841 ATGGCTCTGACGGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATG 900
DB 841 ATGGCTCTGACGGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATG 900
QY 901 TTACGGGATCGGCCAAAAGAGCGCTTGCGATCTCGCTTCGGCGGATCCGGAGCGG 960
DB 901 TTACGGGATCGGCCAAAAGAGCGCTTGCGATCTCGCTTCGGCGGATCCGGAGCGG 960
QY 961 TGATTGAAGCGATCGAGATGAGGACGAGCAATTTCTGTCANAATTTCCCGCGCTAGCA 1020
DB 961 TGATTGAAGCGATCGAGATGAGGACGAGCAATTTCTGTCANAATTTCCCGCGCTAGCA 1020
QY 1021 AAAAAACGGCAAGGCAGATCATCTCTTGACCTGAAA 1055
DB 1021 AAAAAACGGCAAGGCAGATCATCTCTTGACCTGAAA 1055
```

RESULT 2  
AX954657  
LOCUS  
DEFINITION Sequence 123 from Patent WO03093453.  
AX954657 1522 bp DNA linear PAT 08-JAN-2004

ACCESSION AX954657  
VERSION AX954657.1  
KEYWORDS GI:40784029  
SOURCE Bacillus licheniformis  
ORGANISM Bacillus licheniformis  
REFERENCE 1 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
AUTHORS Andersen, J.T., j Rgensen, S.T., Rasmussen, M.D., Olsen, P.B. and Clausen, I.G.  
TITLE Improved bacillus host cell  
JOURNAL Patent: WO 03093453-A 123 13-NOV-2003;  
Novozymes A/S (DK)  
FEATURES Location/Qualifiers  
source 1. 1522  
/organism="Bacillus licheniformis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1402"  
501..1022  
/notes="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF05523.1"  
/db\_xref="GI:40784030"  
DGVGIENKVEAARTLEDFKAAVKGMOLIDQKGFILFRKQVDIDISPLSKNGYIGVT  
EDGVISTPHCRPGILSEPIQSFPQIDIKRLESRMADDLRKIGIPYRTKKEPHEIVAK  
SSGSQHHVEDMKT"

## ORIGIN

Query Match 100.0%; Score 1055; DB 6; Length 1522;  
Best Local Similarity 100.0%; Pred. No. 6.7e-279;  
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTTAAGGTAAAGCAAAAAGAGGTGATATTGATGTACAGCCGAGCAAGATTCAAAATC 60  
DB 468 CCTTAAGGTAAAGCAAAAAGAGGTGATATTGATGTACAGCCGAGCAAGATTCAAAATC 527  
QY 61 GGTATTATGCTTATTGGAAGTCTGCTGCGCGCTCAGCTTTCACCTGGAGCGCTTGCGC 120  
DB 528 GGTATTATGCTTATTGGAAGTCTGCTGCGCGCTCAGCTTTCACCTGGAGCGCTTGCGC 587  
QY 121 GAAAGCGCGCTAAAGTTCAAAATCCAGCTTGAAAGGTTTATCTGGACGGAGACGTTGGA 180  
DB 588 GAAAGCGCGCTAAAGTTCAAAATCCAGCTTGAAAGGTTTATCTGGACGGAGACGTTGGA 647  
QY 181 ATTGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGACTTTTAAAGTGTCTTATAAAGGG 240  
DB 648 ATTGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGACTTTTAAAGTGTCTTATAAAGGG 707  
QY 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTCGACGACATT 300  
DB 708 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTCGACGACATT 767  
QY 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 360  
DB 768 TCTCCCTCAGCAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 827  
QY 361 TTTCACGCTCGCGGGCATCTTATCAGAAACCAATTCATCGTTTTCAGATTGATATA 420  
DB 828 TTTCACGCTCGCGGGCATCTTATCAGAAACCAATTCATCGTTTTCAGATTGATATA 887  
QY 421 AAGCGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAAA 480  
DB 888 AAGCGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAAA 947  
QY 481 AAGGAATTGAACATGTCATTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 540  
DB 948 AAGGAATTGAACATGTCATTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 1007  
QY 541 GATATGAAGACATGACGCTGTTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 600  
DB 1008 GATATGAAGACATGACGCTGTTATGCTCTTTTTCAGCTGCGACAGAGAGCTTTTTCAGCG 1067

```
QY 601 AACATATGTTAACTTTTTCATCTCTAGCTTGGCTGCTTTTGTGTACAAATGAAGACGATC 660
Db 1068 AACATATGTTAACTTTTTCATCTCTAGCTTGGCTGCTTTTGTGTACAAATGAAGACGATC 1127
QY 661 AAAGAGGTGAATGAACGTTTGATCGAATTCGTAAGAGGACGATTCGATTTATGTATCGCCCC 720
Db 1128 AAAGAGGTGAATGAACGTTTGATCGAATTCGTAAGAGGACGATTCGATTTATGTATCGCCCC 1187
QY 721 AATATATTGTCAATGAAAAACGCGGGATCGGCTATFCAGATCTTACGCGCAAAATCGGTTTA 780
Db 1188 AATATATTGTCAATGAAAAACGCGGGATCGGCTATFCAGATCTTACGCGCAAAATCGGTTTA 1247
QY 781 TTTATAAGAAAAACGAAAGCAAACTATACATACATTTATGTATAGAGAACACGA 840
Db 1248 TTTATAAGAAAAACGAAAGCAAACTATACATACATTTATGTATAGAGAACACGA 1307
QY 841 ATGCGCTGTACGGCTTTTCGAAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATG 900
Db 1308 ATGCGCTGTACGGCTTTTCGAAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATG 1367
QY 901 TTACGGGGATCGGCCCAAAAGAGCGCTTTCGATCTCTCGCTTCCGGCGATCCGGAGCGG 960
Db 1368 TTACGGGGATCGGCCCAAAAGAGCGCTTTCGATCTCTCGCTTCCGGCGATCCGGAGCGG 1427
QY 961 TGATTGAAGCGATCGAGAAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGGTAGGCA 1020
Db 1428 TGATTGAAGCGATCGAGAAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGGTAGGCA 1487
QY 1021 AAAAAGCGCAAGGACGATCATCTTTGACCTGAAA 1055
Db 1488 AAAAAGCGCAAGGACGATCATCTTTGACCTGAAA 1522
```

RESULT 3

AE017333\_27/c

WPCOMMENT

Sequence split into 43 fragments LOCUS AE017333 Accession AE017333

Fragment Name	Begin	End
AE017333_00	1	110000
AE017333_01	100001	210000
AE017333_02	200001	310000
AE017333_03	300001	410000
AE017333_04	400001	510000
AE017333_05	500001	610000
AE017333_06	600001	710000
AE017333_07	700001	810000
AE017333_08	800001	910000
AE017333_09	900001	1010000
AE017333_10	1000001	1110000
AE017333_11	1100001	1210000
AE017333_12	1200001	1310000
AE017333_13	1300001	1410000
AE017333_14	1400001	1510000
AE017333_15	1500001	1610000
AE017333_16	1600001	1710000
AE017333_17	1700001	1810000
AE017333_18	1800001	1910000
AE017333_19	1900001	2010000
AE017333_20	2000001	2110000
AE017333_21	2100001	2210000
AE017333_22	2200001	2310000
AE017333_23	2300001	2410000
AE017333_24	2400001	2510000
AE017333_25	2500001	2610000
AE017333_26	2600001	2710000
AE017333_27	2700001	2810000
AE017333_28	2800001	2910000
AE017333_29	2900001	3010000
AE017333_30	3000001	3110000
AE017333_31	3100001	3210000
AE017333_32	3200001	3310000
AE017333_33	3300001	3410000
AE017333_34	3400001	3510000

AE017333_35	3500001	3610000
AE017333_36	3600001	3710000
AE017333_37	3700001	3810000
AE017333_38	3800001	3910000
AE017333_39	3900001	4010000
AE017333_40	4000001	4110000
AE017333_41	4100001	4210000
AE017333_42	4200001	4222645
Continuation (28 of 43) of AE017333 from base 2700001 (AE017333 Bacillus licheniformis D		
Query Match 99.3%; Score 1047.4; DB 1; Length 110000;		
Best Local Similarity 99.9%; Pred. No. 1.le-276;		
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY 7	GGTAAGGCAAAAAAGAGGTGATATTGATGTACAGCCGAAGCAAGTTCAAAAATCGGTTTA	66
Db 92432	GTTAAGGCANAAAGAGGTGATATTGATGTACAGCCGAAGCAAGTTCAAAAATCGGTTTA	92373
QY 67	TTGCTTATTGGAAGTCTGCTGGCCGCTCAGCTTTTCACTGGAGGCCTTGGCCGAAAAG	126
Db 92372	TTGCTTATTGGAAGTCTGCTGGCCGCTCAGCTTTTCACTGGAGGCCTTGGCCGAAAAG	92313
QY 127	CCGGCTAAAGTTCAAAATCCAGCTTGAAGAGGTTTATCTGGACGGAGAGCTTGGAAATCG	186
Db 92312	CCGGCTAAAGTTCAAAATCCAGCTTGAAGAGGTTTATCTGGACGGAGAGCTTGGAAATCG	92253
QY 187	AATAAGTAGAGGCGGCTCGCACACTGGAAGACTTTAAAGCTTTAAAGCTGCTTATAAAGGGTGGCAG	246
Db 92252	AATAAGTAGAGGCGGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAAGGGTGGCAG	92193
QY 247	CTCATTCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAAAACAGGTGGAGCATTTTCTCCC	306
Db 92192	CTCATTCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAAAACAGGTGGAGCATTTTCTCCC	92133
QY 307	CTCAGCAAAACAAACGGTTATATCGAGTGTACTGAGAGTGGCGTATTTTCGATTTTTCAC	366
Db 92132	CTCAGCAAAACAAACGGTTATATCGAGTGTACTGAGAGTGGCGTATTTTCGATTTTTCAC	92073
QY 367	GGTCCCGCGGCATCTTATCAGAAACCCATTCAAATCGTGTGTTTTCAGATTGATATAAGCGG	426
Db 92072	GGTCCCGCGGCATCTTATCAGAAACCCATTCAAATCGTGTGTTTTCAGATTGATATAAGCGG	92013
QY 427	CTGAAAACCGCGATGGCGGATGATCTGCGCAAAAGGATACCATCCGACGAAAGAA	486
Db 92012	CTGAAAACCGCGATGGCGGATGATCTGCGCAAAAGGATACCATCCGACGAAAGAA	91953
QY 487	TTTCGAACATGTCTATTGAAGCCGTAAATCATCCGGAAGCCAAACATCATGTAGAAGATATG	546
Db 91952	TTTCGAACATGTCTATTGAAGCCGTAAATCATCCGGAAGCCAAACATCATGTAGAAGATATG	91893
QY 547	AAGACATGACGCTGTTATGCTTTTTCAGCTGCGAGCAGAGCTTTTTCAGCGAACATA	606
Db 91892	AAGACATGACGCTGTTATGCTTTTTCAGCTGCGAGCAGAGCTTTTTCAGCGAACATA	91833
QY 607	TGTTAACTTTTTCATCTAGCTTTGCTGTTTGTGTGTACAAATGAAGACGAGTCAAAGAG	666
Db 91832	TGTTAACTTTTTCATCTAGCTTTGCTGTTTGTGTGTACAAATGAAGACGAGTCAAAGAG	91773
QY 667	GTGAATGAACGTTGATCGAATTCGTAAAGGGAGCATGATGATGATCGCCCAATATA	726
Db 91772	GTGAATGAACGTTGATCGAATTCGTAAAGGGAGCATGATGATGATGATCGCCCAATATA	91713
QY 727	TTGTTCATTTGAAAACGGCGGATCGGCTATCAGATCTTTCACGCCAAATCCGTTTATTATA	786
Db 91712	TTGTTCATTTGAAAACGGCGGATCGGCTATCAGATCTTTCACGCCAAATCCGTTTATTATA	91653
QY 787	AGAAAAACAGCAAAAGAAACAAATCTATACATACCATTTATGTAAAGAGAACACGATGCGC	846
Db 91652	AGAAAAACAGCAAAAGAAACAAATCTATACATACCATTTATGTAAAGAGAACACGATGCGC	91593
QY 847	TGTACGGCTTTTCGACAAAGGGAAGAAAAATGCTGTTTTCGAAAAATGCTGAATGTTACGG	906
Db 91592	TGTACGGCTTTTCGACAAAGGGAAGAAAAATGCTGTTTTCGAAAAATGCTGAATGTTACGG	91533

Qy 907 GGATCGGCCCAAGAGCGCTTGGATCTCGCTTCGGCGATCCGGAGCGGTGATTG 966  
|||||  
Db 91532 GGATCGGCCCAAGAGCGCTTGGATCTCGCTTCGGCGATCCGGAGCGGTGATTG 91473  
|||||  
Qy 967 AAGCGATCGAAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGGGGTAGGCAAAAAA 1026  
|||||  
Db 91472 AAGCGATCGAAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGGGGTAGGCAAAAAA 91413  
|||||  
Qy 1027 CGGCAAGCGAGATCATCTTGGACCTGAAA 1055  
|||||  
Db 91412 CGGCAAGCGAGATCATCTTGGACCTGAAA 91384  
|||||

RESULT 4  
CP000002\_27/c  
WPCOMMENT  
Sequence split into 43 fragments LOCUS CP000002 Accession CP000002  
Fragment Name Begin End

CP000002\_00 1 110000  
CP000002\_01 100001 210000  
CP000002\_02 200001 310000  
CP000002\_03 300001 410000  
CP000002\_04 400001 510000  
CP000002\_05 500001 610000  
CP000002\_06 600001 710000  
CP000002\_07 700001 810000  
CP000002\_08 800001 910000  
CP000002\_09 900001 1010000  
CP000002\_10 1000001 1110000  
CP000002\_11 1100001 1210000  
CP000002\_12 1200001 1310000  
CP000002\_13 1300001 1410000  
CP000002\_14 1400001 1510000  
CP000002\_15 1500001 1610000  
CP000002\_16 1600001 1710000  
CP000002\_17 1700001 1810000  
CP000002\_18 1800001 1910000  
CP000002\_19 1900001 2010000  
CP000002\_20 2000001 2110000  
CP000002\_21 2100001 2210000  
CP000002\_22 2200001 2310000  
CP000002\_23 2300001 2410000  
CP000002\_24 2400001 2510000  
CP000002\_25 2500001 2610000  
CP000002\_26 2600001 2710000  
CP000002\_27 2700001 2810000  
CP000002\_28 2800001 2910000  
CP000002\_29 2900001 3010000  
CP000002\_30 3000001 3110000  
CP000002\_31 3100001 3210000  
CP000002\_32 3200001 3310000  
CP000002\_33 3300001 3410000  
CP000002\_34 3400001 3510000  
CP000002\_35 3500001 3610000  
CP000002\_36 3600001 3710000  
CP000002\_37 3700001 3810000  
CP000002\_38 3800001 3910000  
CP000002\_39 3900001 4010000  
CP000002\_40 4000001 4110000  
CP000002\_41 4100001 4210000  
CP000002\_42 4200001 4222334

Continuation (28 of 43) of CP000002 from base 2700001 (CP000002 Bacillus licheniformis A

Query Match 99.3%; Score 1047.4; DB 1; Length 110000;  
Best Local Similarity 99.9%; Pred. No. 1.1e-276;  
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 7 GGTAAAGCAAAAGAGGTATATTGATGTACAGCCGCAAGCAAGTTCAAATCGGTTA 66  
|||||  
Db 91570 GTTAAAGCAAAAGAGGTATATTGATGTACAGCCGCAAGCAAGTTCAAATCGGTTA 91511  
|||||  
Qy 67 TTGCTTATTGGAAGTCTGCTGCCGCGCTCAGCTTTACCTGGAGGCGCTTGGCCGAAAAAG 126

Db 91510 TTGCTTATTGGAAGTCTGCTGCCGCGCTCAGCTTTACCTGGAGGCTTGGCCGAAAAAG 91451  
|||||  
Qy 127 CCGGCTAAAGTTCAAAATCCAGCTTGAAAGGTTTATCTGACCGAGAGCGTTGGAAATGAG 186  
|||||  
Db 91450 CCGGCTAAAGTTCAAAATCCAGCTTGAAAGGTTTATCTGACCGAGAGCGTTGGAAATGAG 91391  
|||||  
Qy 187 AATAAGTAGAGCCGCTCCACACTGGAAGACTTTTAAAGCTGCTTAAAGSGTGGCAG 246  
|||||  
Db 91390 AATAAGTAGAGCCGCTCCACACTGGAAGACTTTTAAAGCTGCTTAAAGSGTGGCAG 91331  
|||||  
Qy 247 CTCAATCGATCAGAAAAAGGGGTTTATCTGTTTCGCAACAGGTGGAGCATTTCTCCCC 306  
|||||  
Db 91330 CTCAATCGATCAGAAAAAGGGGTTTATCTGTTTCGCAACAGGTGGAGCATTTCTCCCC 91271  
|||||  
Qy 307 CTCAGCAAAAACAAACGGTTATATCGGAGTGAAGTGAAGTGGCGTGAATTCACATTTTCCAC 366  
|||||  
Db 91270 CTCAGCAAAAACAAACGGTTATATCGGAGTGAAGTGAAGTGGCGTGAATTCACATTTTCCAC 91211  
|||||  
Qy 367 GGTCCGCCGGGCATCTTATCAGAACCCATTCNAATCGTTTTTTCAGATTGTATAAAGCGG 426  
|||||  
Db 91210 GGTCCGCCGGGCATCTTATCAGAACCCATTCNAATCGTTTTTTCAGATTGTATAAAGCGG 91151  
|||||  
Qy 427 CTGAAAAGCCGGATGGCGGATGATCTCGCAAAAGGGATACCATACCGCACGAAAAAGGAA 486  
|||||  
Db 91150 CTGAAAAGCCGGATGGCGGATGATCTCGCAAAAGGGATACCATACCGCACGAAAAAGGAA 91091  
|||||  
Qy 487 TTGGAACATGTCATTGAAGCGGTAAATCATCCGGAAGCCAAATCATGTAGAGATATG 546  
|||||  
Db 91090 TTGGAACATGTCATTGAAGCGGTAAATCATCCGGAAGCCAAATCATGTAGAGATATG 91031  
|||||  
Qy 547 AAGACATGACGCTGTTATGCTTTTTTTCAGCTCAGACAGAGCTTTTTTAGCGAAACATA 606  
|||||  
Db 91030 AAGACATGACGCTGTTATGCTTTTTTTCAGCTCAGACAGAGCTTTTTTAGCGAAACATA 90971  
|||||  
Qy 607 TGTTAACCTTTTCACTTAGCTTTGCTGTTTGTGTTTCAATGAAGAGCAGTCAAAAGAG 666  
|||||  
Db 90970 TGTTAACCTTTTCACTTAGCTTTGCTGTTTGTGTTTCAATGAAGAGCAGTCAAAAGAG 90911  
|||||  
Qy 667 GTGAATGAACGTTGATCGAATTCGTAAGAGGAGATGATTTATGTCGCCCCCAATATA 726  
|||||  
Db 90910 GTGAATGAACGTTGATCGAATTCGTAAGAGGAGATGATTTATGTCGCCCCCAATATA 90851  
|||||  
Qy 727 TTGTCATTGAAAAGCGGGGATCGGCTATCAGATCTTTCAGCCAAATCCGTTTATTATA 786  
|||||  
Db 90850 TTGTCATTGAAAAGCGGGGATCGGCTATCAGATCTTTCAGCCAAATCCGTTTATTATA 90791  
|||||  
Qy 787 AGAAAAACAGCAAGAAACAACTTATATACATACCACTTATGTAAGAGAAAGACACGAATGCGC 846  
|||||  
Db 90790 AGAAAAACAGCAAGAAACAACTTATATACATACCACTTATGTAAGAGAAAGACACGAATGCGC 90731  
|||||  
Qy 847 TGTAAGGCTTTTTCGCAAGGGAAGAAAAATGCTGTTTTCGAAAAATGCTGAATGTTACGG 906  
|||||  
Db 90730 TGTAAGGCTTTTTCGCAAGGGAAGAAAAATGCTGTTTTCGAAAAATGCTGAATGTTACGG 90671  
|||||  
Qy 907 GGTCCGCCCAAAAGAGGCGCTTCCGATCTCCGCTCCGCGATCCGCGAGCGGTGATTG 966  
|||||  
Db 90670 GGTCCGCCCAAAAGAGGCGCTTCCGATCTCCGCTCCGCGATCCGCGAGCGGTGATTG 90611  
|||||  
Qy 967 AAGCGATCGAAGATGAGGAGCAAGCATTTCTCGTCAAAATTTCCCGGGGTAGGCAAAAAA 1026  
|||||  
Db 90610 AAGCGATCGAAGATGAGGAGCAAGCATTTCTCGTCAAAATTTCCCGGGGTAGGCAAAAAA 90551  
|||||

RESULT 5  
AX434675  
LOCUS  
DEFINITION Sequence 3090 from Patent WO0229113.  
ACCESSION AX434675  
AX434675 459 bp DNA linear PAT 28-JUN-2002

```

VERSION      AX434675.1  GI:21659483
KEYWORDS
SOURCE       Bacillus licheniformis
ORGANISM     Bacillus licheniformis
REFERENCE    1
AUTHORS      Berka, R. and Clausen, I. G.
TITLE        Methods for monitoring multiple gene expression
JOURNAL      Patent: WO 0229113-A 3090 11-APR-2002;
              Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES     Location/Qualifiers
source       1. 459
              /organism="Bacillus licheniformis"
              /mol_type="unassigned DNA"
              /db_xref="taxon:1402"
ORIGIN
Query Match      43 5%; Score 459; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 7.1e-115; Mismatches 0; Indels 0; Gaps 0;
Matches 459; Conservative 0;

Qy 65 TATTCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTTCACTGGAGGCTTGGCCGAAA 124
Db 1 TATTGCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTTCACTGGAGGCTTGGCCGAAA 60

Qy 125 AGCCGGCTAAAGTTCAAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 184
Db 61 AGCCGGCTAAAGTTCAAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 120

Qy 185 AGAATAAGTAGAGCGCCCTGCACACTGGAAGACTTTAAAGCTGCTTATAAGGGTGCC 244
Db 121 AGAATAAGTAGAGCGCCCTGCACACTGGAAGACTTTAAAGCTGCTTATAAGGGTGCC 180

Qy 245 AGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAAAACAGGTGGACGACATTTCCTC 304
Db 181 AGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAAAACAGGTGGACGACATTTCCTC 240

Qy 305 CCCTCAGCAAAACAAACCGTTATATCGAGTGACTGAAGATGGCGTGAATTCGACTTTTC 364
Db 241 CCCTCAGCAAAACAAACCGTTATATCGAGTGACTGAAGATGGCGTGAATTCGACTTTTC 300

Qy 365 AGCGTCGCCCGGCATCTTATCGAACCCATTCATCTGTTTTCAGATTCATATAAGC 424
Db 301 AGCGTCGCCCGGCATCTTATCGAACCCATTCATCTGTTTTCAGATTCATATAAGC 360

Qy 425 GCCTGGAAGCCGGATGCGGATGATCTGCGCAAAAGGATACCATACCGCAGCAAAAAGG 484
Db 361 GCCTGGAAGCCGGATGCGGATGATCTGCGCAAAAGGATACCATACCGCAGCAAAAAGG 420

Qy 485 AATTTGAACATGCTATTGAAGCCGTAAATATCATCCGGAA 523
Db 421 AATTTGAACATGCTATTGAAGCCGTAAATATCATCCGGAA 459

RESULT 6
BSY15896
LOCUS
DEFINITION  BSY15896 17787 bp DNA linear BCT 18-APR-2005
            Bacillus subtilis nadA, yrbA, yrbB, yrbC, yrbD, orf7, yrbE, csbX,
            bofC, ruvA, ruvB, queA, tgt, yrbF, orf16, yrbG, spoVB genes and
            partial nadC gene.
ACCESSION  Y15896
VERSION    Y15896.1  GI:6977794
KEYWORDS   bofC gene; csbX gene; nadA gene; nadC gene; nicotinate-nucleotide
            pyrophosphorylase; queA gene; quinolinolate synthetase; ruvA gene;
            yrbB gene; sporulation-specific protein; spoVB gene; tgt gene;
            tRNA-guanine transglycosylase; yrbA gene; yrbB gene; yrbC gene;
            yrbD gene; yrbE gene; yrbF gene; yrbG gene.
SOURCE     Bacillus subtilis
ORGANISM   Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE  1
AUTHORS    Tosato, V., Bolotin, A., Bertani, I., Valentino, I. and Bruschi, C. V.
TITLE      A 17.8 kb segment in the spoVB-nadC region of the Bacillus subtilis
```

```

168 chromosome: sequencing and ruv operon identification
Unpublished
REFERENCE    2 (bases 1 to 17787)
AUTHORS     Bruschi, C. V.
TITLE       Direct Submission
JOURNAL     Submitted (11-DEC-1997) C.V. Bruschi, International Centre for
            Genetic Engineering and Biotechnology, Padriciano 99, I-34012
            Trieste, ITALY
FEATURES     Location/Qualifiers
source       1. 17787
              /organism="Bacillus subtilis"
              /mol_type="genomic DNA"
              /strain="168"
              /db_xref="taxon:1423"
              /map="242-243"
              <1. .167
              /gene="nadC"
              <1. .167
              /gene="nadC"
              /function="NAD biosynthesis"
              /codon_start=3
              /transl_table=11
              /product="nicotinate-nucleotide pyrophosphorylase"
              /protein_id="CAB75320.1"
              /db_xref="GI:6977795"
              /db_xref="GOA:Q9KWZ2"
              /db_xref="InterPro:IPR002638"
              /db_xref="UniProt/TREMBL:Q9KWZ2"
              /translation="ASGGITLESIPAFKGTGVNYSIGFLTHSVKSLDISMDVTLSNE
              SVECCTVNS"
              169. .1260
              /gene="nadA"
              169. .1260
              /gene="nadA"
              /function="NAD biosynthesis"
              /codon_start=1
              /transl_table=11
              /product="quinolinolate synthetase A"
              /protein_id="CAB75321.1"
              /db_xref="GI:6977796"
              /db_xref="GOA:Q9KWZ1"
              /db_xref="UniProt/Swiss-Prot:Q9KWZ1"
              /translation="MIKQSDNMPSEYKELSRKDMETRVAAIKKFGSRLFI PGHHYQ
              KDEVIQADQDGTDSLQAOVAKNKEADYIVFCGVHPMAETADMLTSSQQTWVLPMR
              AGCSMDADMQQTNRAWKQLQHI FGDTIIPLYVNSTABI KAFVKGHGATVTSNA
              KVLWFAPTOKKRILFLPDQHLGRNTAVDLGIALEDMAVWDPKMDLVAESGHTNVKV
              ILWKGHCSVHEKPTTKNIHDMREDPDIOIIVHPECSEHVVTLSDDNGSTKYIIDTIN
              QAPAGSKWATGEMNLVQRIIHEHPDKQIESLNPDMPCLTMNRIDILPHLLWSLEQIE
              KGPESGVIKVPKAIQEDALLALNRLSIT"
              1377. .2540
              /gene="yrbA"
              1377. .2540
              /gene="yrbA"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
              /protein_id="CAB75322.1"
              /db_xref="GI:6977797"
              /db_xref="GOA:O32062"
              /db_xref="InterPro:IPR002482"
              /db_xref="UniProt/TREMBL:O32062"
              /translation="MKIHIQGDLSLWKIAEKYGVDEVVKLNTQLSNPDLIMPGMK
              IKVPSGVPRKPKAGSVPAGSVKQEHYPYKPKSVVDVDTTKPKKKSMYPVPP
              MPNIQENVYADVNDYDMKQLFQWSPKPKPKKHGDMHDMHYHMQDQFPQOEAE
              MSNMENYNNPNPKAPKPEVGGIIEENHHTVNMMPAPQPYTHYPAHFVPCFPV
              SPILPGLGCLCYPPYPAQYPMHPMHGYPQVSPQYDPGYENQHSHHHGHYGSYGA
              PQASPAYSPYGHMPYPYGTTPQVMGAYQPAARAHYMPYKDHDDCCDGDHQPFYS
              APGSHSGMAYGSPNNPYGTANPNPNPYSGVSMPTNQPSVNMFGPRESENE"
              2766. .3284
              /gene="yrbB"
              2766. .3284
              /gene="yrbB"
              /codon_start=1
```

```

/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75323.1"
/db_xref="GI:6977798"
/db_xref="GOA:P94446"
/db_xref="UniProt/Swiss-Prot:P94446"
/transl_table="MNTTRNNGTTPGTYTNNENDADRQGDGIDHDGPFVSELMEDQND
GNRTTNNNRDRVTADDRVPLATDGTNNNNMDNRNANNQYDQENRRLLAAKTA
NRVKQNVANDQVMVSDRRVVIKSHREFTKSDRNVVKAARYANGRDVQVSTDK
GLRKLHMMNR"
3387..4109
/gene="yrbC"
3387..4109
/gene="yrbC"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75324.1"
/db_xref="GI:6977799"
/db_xref="GOA:P94447"
/db_xref="InterPro:IPR002876"
/db_xref="UniProt/Swiss-Prot:P94447"
/transl_table="MAGHSKWKNIQKRNQADAKRKI FMKLAKEIYVAKEGGADPE
ANSALRIEKAAGAMPNENIDRAIKKAAGQDGSSEYETTYEGYGPSGIYAVMECV
TDNKRITASNVRTAFNKGSLGSGCVAFLEFKRGFTTIDRTERQIEDEELMDVLE
AGGELRIEEDLYXEVTEPEHFEVETALBSKYLIISSAEVTMLPNTYAEADDQAVEKL
ETLLDVELEDDDDQEVVTNRS"
complement(4150..5604)
/gene="yrbD"
complement(4150..5604)
/gene="yrbD"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75325.1"
/db_xref="GI:6977800"
/db_xref="GOA:O32060"
/db_xref="InterPro:IPR001463"
/db_xref="InterPro:IPR002293"
/db_xref="UniProt/TrEMBL:O32060"
/transl_table="MADFVSLNVLMTSTPIYILLGIGFAPFSIMTRFLQVRHLKEMI
VQMFKGSSEAGVSFALSIALSRVGTGNIAGVATAIAFGGFGAVFWMAIAFIGA
ASAPVESTLAQIVKVGQYRGGRVYIEKGLIKWFAVLFAAAALIAMAFILMPGVQ
SNSAAGIONAFGLSPFVTCGLVLLGFIIFGVKRIANNAOMIYVPMALGYLLSL
IIIIWVSELPAVLSLIFKSAFALDSAFGLIGMAISGWKRGYISNEAGCGTGPHPA
AAAEVRPQGLVQARSVYIDTLFVCSATAFMLFTGMVNTQADGSGFIVHQLKVE
AGPQFTQAIDSVLPGFGAGEVAIALFFFAFTIMYYTYAETNIAYIARGRESKWM
LGLKLIILAAFTYGTVKTASLAWALGDAGLGMVNLVNIATVLLAKPALLAKDYERQ
KKOGLDPIFDPKALGIKNADFEKEYTHESERVS"
complement(6712..7104)
/note="orf7"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75326.1"
/db_xref="GI:6977801"
/db_xref="GOA:O32057"
/db_xref="InterPro:IPR000719"
/db_xref="UniProt/TrEMBL:O32057"
/transl_table="MLTKAEALAHTVIYRKNRFRDKVKEKSEELTLIGKRSAYVFL
TEGGRKVALKVFPEYQATAVKEAAIYKLAGSASYPDIYETGDSFILMEYIKETFY
NCLKKGTASIDDMIQQVEEALSDAQPD"
7342..8367
/gene="yrbE"
7342..8367
/gene="yrbE"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75327.1"
/db_xref="GI:6977802"
/db_xref="GOA:O05389"
/db_xref="InterPro:IPR000683"

```

gene

CDS

gene

CDS

CDS

gene

CDS

gene  
CDSgene  
CDSQuery Match  
Best Local Similarity 43.3%; Score 457; DB 1; Length 17787;  
Matches 659; Conservative 0; Mismatches 285; Indels 11; Gaps 2;

```

Qy 101 TTCACTCTGGAGCGCTTGGCCGAAAAGCGGCTAAAGTTCAAATCCAGCTTCAAAAGGTTT 160
Db 9878 TTGGCGCGGAAGTGGACACATATGAGCCGCTGCGAGTGCATGTCGACCTTGAAGAAGTGT 9937
Qy 161 ATCTGGAAGGAGCGTGGAAATGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGACT 220
Db 9938 ATTTGGACGCGGATGTGAGCATTCAGCATTAAGCATGAAAGATGTTTCTATGATGATT 9997
Qy 221 TTAAGCTGCTTATAAGGGTGGCAGCTCATCGATCAGAAAAGGGGTTTATCTGTTTC 280
Db 9998 TTTGGCGAGCTTATGCGGATGGACGCTAGTGGAAACAAAAGGGATATGTCCTCTCC 10057
Qy 281 GCAACAGGTGGACGACATTTCTCCCTCAGCAAAAACAAACGGTTATATCGAGTGACTG 340
Db 10058 GGAAGCAATGGACGATATTTCTCCGCTCAGCAGGTCAACGGTATATCGGTATCGG 10117
Qy 341 AAGATGGCGTGAATTCGACTTTTTCACGGTCCCGCGGCATCTTATCAGAAACCAATCAAT 400
Db 10118 ATAATGAGTGATTTCCACTTTTTCATGGCGGCGCTGAGCCAGCTTCCGAACCAATTCAGT 10177
Qy 401 CGTTTTTTCAGATTGATATAAGCGGCTGGAAAGCCGATGCGGATGATCTGCGCAAG 460
Db 10178 CTTTTTTTCAAAATGATTTAGAAAGGCTGGAAAGCCATATGCAAAAAAATCTGCTGAAAG 10237
Qy 461 GGATACCATACCGCACGAAAAGGAATTTGAACATGTCATTGAAGCCGTAATAATCATCCG 520
Db 10238 GCATTCATTTCCGACAAAAGCGAGTTTGAGGATGTCATCGAA-----CATATG 10287
Qy 521 GAAGCCCAACATCATGTATAGAGATATGAAGACATGACCGCTGTATGTCTTTTTCAGCTGC 580
Db 10288 AAGACATACAGCGGTAGCGCTCGCTGATTGAAGACATGATTATTCATGCTCTTTTTCGT 10347
```









QY	881	GTATTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAGAGCGCTTGCGATCCTCGC	940
Db	22993	CTTTACAAAGCTGTTAAACGTCGACCGAATTCGCCGAAAGGAGCACCTGCCATCTTGGG	22934
QY	941	TTCCGGCGATCGGAGCGGTGATTGAACCGATCGAGATGAGGACGAATTTCTCGT	1000
Db	22933	CTCGGGTATCCCGAGCGGTGATCCAAAGCAATCGAAATGAAGATGAAGCTTTTTTGGT	22874
QY	1001	CAAAATTTCCCGCGTAGGCAAAAAACGCGAAGGAGATCATCTCTTGACCTGAAA	1055
Db	22873	CAAGTTTCTCGAGTAGGAAGAAACCGCAGCCAGATTAATCTAGATTAAAA	22819
RESULT 8			
BSBOFCGEN			
LOCUS	BSBOFCGEN	2664 bp DNA linear	BCT 18-APR-2005
DEFINITION	B. subtilis bofC, orf1, csbX, and orf4 genes.		
ACCESSION	X93081		
VERSION	X93081.1	GI:1941915	
KEYWORDS	bofC gene; csbX gene; ORF1; ORF4.		
SOURCE	Bacillus subtilis		
ORGANISM	Bacillus subtilis		
REFERENCE	1	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
AUTHORS	Gomez,M. and Cutting,S.M.		
TITLE	BofC encodes a putative forespore regulator of the Bacillus		
JOURNAL	subtilis sigma K checkpoint		
PUBMED	Microbiology (Reading, Engl.) 143 (PT 1), 157-170 (1997)		
REFERENCE	9025289		
AUTHORS	2 (bases 1 to 2664)		
TITLE	Cutting,S.M.		
JOURNAL	Direct Submission		
FEATURES	Submitted (14-NOV-1995) S.M. Cutting, Dept. of Microbiology, University of Pennsylvania School of Medicine, 345 Johnson Pavillion, 3610 Hamilton Walk, Philadelphia, PA 19104-6076, USA		
source	Location/Qualifiers		
	1. .2664		
	/organism="Bacillus subtilis"		
	/mol_type="genomic DNA"		
	/strain="PV79"		
	/isolate="168"		
	/db_xref="taxon:1423"		
gene	/germline		
	<1. .275		
	/gene="orf1"		
	<1. .275		
	/gene="orf1"		
	/codon_start=3		
	/transl_table=11		
	/protein_id="CAA63619.1"		
	/db_xref="GI:1941916"		
	/db_xref="GOA:O05389"		
	/db_xref="UniProt/Swiss-Prot:O05389"		
gene	/translation="AAADNRPTTVEVSTADFVMDKPKHFFFLERYKDSYBEEILRFA EAIGNQETPTCGNDGLQAGRIARAAQOGLAFGMPVSIETKEIAF"		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAA63620.1"		
	/db_xref="GI:1941917"		
	/db_xref="GOA:O05390"		
	/db_xref="UniProt/Swiss-Prot:O05390"		
gene	/translation="MNTVHAKGNVLNKIGIPSHMWVGIVGVIWVGDLGQWLSFP LVDHLSMQQSASLFTMYGIATVTSAMLSTGTVETWGRKMTVGLLAFILGSAFIG WAIPIHWYPALLGSLYALRGYLPFLAYSFLVWSYSTSONILKXAVGFWFMPTCGLN VLGPFYSYAPFAGEINLWSALLFVAAGGILALFFNKDKFTPIQKQDPKWLKSR AFTIMFNPKRIVIGVVKTNALIGQFGFAIFLPTLYLARYGYSVSEWLQIWGLTFVNI VFNIIFGAVGDKLAWRNTVMWFGVGCGIFLALYTFQLIQHYYWLMIIAACCCYGAAL		
	346. .1650		
	/gene="csbX"		
	346. .1650		
	/gene="csbX"		
	/note="sigma B transcribed gene"		

Qy	641	TGTTACAATGAAGACAGTCAAGAGGTGAATGAACGTTGATCGAATTCGTAAAGGGAC	700
Db	2313	TGGTACAATGAAGAACAGTCAAGAGGTGAATTTGGC-TGATTGAAATTTGTTAAAGGGAC	2371
Qy	701	GATTGATTATGTATCGGCCCAATATATGTTCATTGGAACCGGGGATCGGTATCAGAT	760
Db	2372	GATTGATTATGTATCTCCCAATATATGTGTCATGAGACGGCGGCATCGGTATCAGAT	2431
Qy	761	CTTCACGCCAAATCCGTTTATTTATAAGAAAAACAGCAAGAAACAATCTATACATACCA	820
Db	2432	TTTCACGCCGAATCCGTTTATATATAAGGAGAGAGCCCAAGAAACGATTTTACATACCA	2491
Qy	821	TTATGTTAAGAGAGACACGAAATGGCGTGTACGGCTTTTTCGACAGGGAAGAAAAATGCT	880
Db	2492	TCATATCAGAGAGGATGCTTTTTCGCTTTTATGGTTTTTCAACAGAGAGAGAGAGCGCT	2551
Qy	881	GTTTACGAAAATGCTGAAATGTACGGGGATCGGCCCAAAAGGAGCGCTTCGATCCTCGC	940
Db	2552	CTTTACAAAGCTGTTAAACGTGACCGGAATTTGGCCGGAAGGAGCATTTGCCATCTTGGG	2611
Qy	941	TTCCGGCGATCCGGGACGGTGTATTGAAGCGATCGAGAATGAGGACGAAGCAT	993
Db	2612	CTCGGTTGATCCCGGAGCGGTCACTCCAAAGCAATTTGAAATTAAGATGAAGCTT	2664

RESIST. 9

RESULI 9 25/0  
BA000043 25/0

BA000043  
WRCOMMENT

COMMENT:  
SOMEONE PUT INTO 26 ELEMENTS 10000 00000000 00000000

Db	108566	A-----ATGATCCGCCATTTTGGCGTCG	108544
Qy	564	TGTCCTTTTTCAGCTGCAGACAGAAGCTTTTTTAGCAACAATATGTTAACTTTTTTCATTC	623
Db	108543	T---CCAAATAAACCAACGAAAGCCCGTGGTGCAAGGCGCATTCGGGGATTTCTTTTTCG	108487
Qy	624	TAGCTTTGCGCTGTTTGTGTACAAATGAAGAGCAGTCAAAGAGGTGAATGAACGTTGATC	683
Db	108486	TATAATGAAATCATATGATACAAATGAAGAGCGAATAAAGAGAGGAGCATCCATTGATC	108427
Qy	684	GAATTCGTAAAGGGAGGATTTGATTATGTCGCCCCCAATATATTTGTCATTGAAACGGC	743
Db	108426	GAGTTTATCCGTGGGTATGTCGATTACGTTCTGCCCGGAATATATCGTCATCGACCATAAC	108367
Qy	744	GGGATCGCTATCAGATCTTCACGCCAAAATCCGTTTATTTATAGAAAAACAGCAA----	799
Db	108366	GGCATCGCTACGAGATCTTTACGCCGAAACCCGTTTTCGGTTTCAGAAAGCCGAGGGG	108307
Qy	800	--AGAAAACAATCTATACATACCATTTATGTAAAGAGAAGACAGCAATGCCTGTACCGCTTT	857
Db	108306	CGCGTGACCGTCTATACATACGATATGTACGCGAAGGTGCGCCCTTGTACGGGTTC	108247
Qy	858	TCGACAAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCA	917
Db	108246	CCAAAGCGCGAGGACGCAACTTCTTCGCCAAGCTGCTGCAAGTGTTCGGGCATCGGGCGG	108187
Qy	918	AAAGGAGCGCTTGCATCCTCGCTTCGGCGATCCGGAGCGGTGATTGAACGATCCGAG	977
Db	108186	AAAGCGCGCTTCGCAATTTTTCGGCGCGCGGCGTCCGATGAACCTGGCGGGGCCATCCGAG	108127
Qy	978	AATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGCTAGGCAAAAAAACGCAAGGCAG	1037
Db	108126	GAAGAAACGAGGCGTTTTTATGCAAGTTTCCAGCGGTGGGCAAAAAAGACGCCCGGCAA	108067
Qy	1038	ATCATCTTGCACCTGAAA	1055
Db	108066	ATGATTTTGGATTTAAAA	108049

RESULT 10  
BA000043 26/c  
WPCOMMENT  
Sequence split into 36 fragments LOCUS BA000043 Accession BA000043

Fragment Name	Begin	End
BA000043_00	1	110000
BA000043_01	100001	210000
BA000043_02	200001	310000
BA000043_03	300001	410000
BA000043_04	400001	510000
BA000043_05	500001	610000
BA000043_06	600001	710000
BA000043_07	700001	810000
BA000043_08	800001	910000
BA000043_09	900001	1010000
BA000043_10	1000001	1110000
BA000043_11	1100001	1210000
BA000043_12	1200001	1310000
BA000043_13	1300001	1410000
BA000043_14	1400001	1510000
BA000043_15	1500001	1610000
BA000043_16	1600001	1710000
BA000043_17	1700001	1810000
BA000043_18	1800001	1910000
BA000043_19	1900001	2010000
BA000043_20	2000001	2110000
BA000043_21	2100001	2210000
BA000043_22	2200001	2310000
BA000043_23	2300001	2410000
BA000043_24	2400001	2510000
BA000043_25	2500001	2610000
BA000043_26	2600001	2710000
BA000043_27	2700001	2810000
BA000043_28	2800001	2910000
BA000043_29	2900001	3010000
BA000043_30	3000001	3110000
BA000043_31	3100001	3210000
BA000043_32	3200001	3310000
BA000043_33	3300001	3410000
BA000043_34	3400001	3510000
BA000043_35	3500001	3547776

Continuation (27 of 36) of BA0000043 from base 2600001 (BA0000043 Geobacillus kaustophilus

Query Match	19.9%;	Score 209.6;	DB 1;	Length 110000;
Best Local Similarity	54.0%;	Pred. No. 4.6e-46;		
Matches	528;	Conservative	0;	Mismatches 404; Indels 46; Gaps 3;
Qy	84	GCTGCCCGCTCAGCTTTCACCTGGAGGCTTGGCCGAAAGCGGCTCAAGTTCAAT	143	
Db	8986	GATGCCGCCATCGTGTTCGCGTTCATTTCGCGCTCCGCGCACCGGTTAAGATGACGAT	8927	
Qy	144	CCAGCTTGAAAGGTTTATCTGGACGAGAGCTTGGAAATGAGATAAAGTAGAGCGCG	203	
Db	8926	TGTGCTGAGCGGCAATCTTGGACGAGAGATGAGCGAAGAAAGTCAACGAAACCGT	8867	
Qy	204	TGCGACACTGGAAGACTTTAAAGCTTTATAAGGTGGCGCTTCATCGATCAGAAAA	263	
Db	8866	CGATTGATGACAGAAATATGGAATAATTCGCGGTGGCAACTCGTACGCTCGATGA	8807	
Qy	264	GGGGTTTATCTGTTTCGCAACAGGTGGAGACATTTCTCCCTCAGCAAAAACAAACGG	323	
Db	8806	TCAGACGATCGTATTTGCAAAAACATCAACGACATTTTCGCGTGTGTTTAAACAAACGG	8747	
Qy	324	TTATATCGGAGTACTGAAGATGGGTGATTTTCGACTTTTCACGGTCCCGCGGATCTT	383	
Db	8746	CTACTTCGGAATATAGGACGATGGCACGTCTGTCTATCTTCAACGCGCAACCGGCGGTC	8687	
Qy	384	ATCAGAACCCATTCAATCGTTTTTTCAGATTGATATAAGCGGTGGAAAGCCGATGGC	443	
Db	8686	GAGCGAATCATTCAGTCTTTTCCAAATCGATGCAAAAGCTTCGAAAGCCGTACGA	8627	
Qy	444	GGATGATCTGCGCAAGGGATACCATACCGCACGAAAGAAATTTGAACATGTCATTGA	503	
Db	8626	AGAAAGCTGAAGAAAGGAATTCGCGTCTCTCGAAGAGCGGTATGAACAGTATCGA	8567	
Qy	504	AGCGTAAATCATCCGGAAGCAACATCATGTAGAGATATGAAGACATGACGCTGTTA	563	

Db	8566	A-----ATGTACCGCCATTTTTCGGTCG	8544	
Qy	564	TGCTCTTTTTCAGCTGCAGACAGAAAGCTTTTTCAGCGAACATATGTTAACTTTTTCATTC	623	
Db	8543	T---CCATAAACCAAGCGAAAGCCGCTGGTGCAGAGCGCATCGCGGATTTCTTTTTCG	8487	
Qy	624	TAGCTTTGCTGTTTGTGTTTACAAATGAAGAGCAGTCAAGAGGTTGAATGAACGTTGATC	683	
Db	8486	TATAATGAATCATATGATACAAATGAAGAGCAGAAAGGAGAGAGGATCCATTGATC	8427	
Qy	684	GAAATCGTAAAGGAGCGATTTGATATGATCGCCCAATATATTTGTCATTTGAAACCGC	743	
Db	8426	GAGTTTATCCGCTGATGTCGTTAATCGTCTGCCCGGAATATATCGTCATCGACCAATAC	8367	
Qy	744	GGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTTATAAGAAAAACAGCAA---	799	
Db	8366	GGCATCGCTACGAGATCTTTACCGCGAACCCGTTTGGCTTTCAGAAAGCGCGAGCG	8307	
Qy	800	--AGAAACAATCTATACATACATTTATGTAAGAGACACGAATGCGCTGTACGGCTTT	857	
Db	8306	CGCGTGACCGTCTATACATACGAATATGTACGCGAAGGTGTGCAGCCTTTGTACGGTTC	8247	
Qy	858	TCGACAAAGGAGAAAAATCTCTTTTACGAAATGCTGAATGTTACGGGATCGGCCCA	917	
Db	8246	CCAAACGCGAGGAGCGCAACTTGTTCGCCAAGCTGTGCAAGTGTTCGGGATCGGGCG	8187	
Qy	918	AAAGAGCGCTTGGATCTCTCGCTTCGGCGATCCGGAGCGGTGATTGAAGCGATCGAG	977	
Db	8186	AAAGCGGCTCGCATTTTGGCGCGCGGCGTTCGGATGAATCTGGCGCGGCGCATCGAG	8127	
Qy	978	AATGAGACGAAAGCATTTCTCGTCAAAATTTCCCGCGGTAGCCAAACCGCAAGCGAG	1037	
Db	8126	GAAGAAAAACGAGCGGTTTATGCAAGTTTCCAGCGTGGGCAAAAGACGCGCCGCA	8067	
Qy	1038	ATCATCTTGCACCTGAAA	1055	
Db	8066	ATGATTGGAATTTAAAA	8049	
RESULT 11				
AX433201	AX433201	546 bp	DNA	linear
LOCUS	Sequence 1616 from Patent WO0229113.			PAT 28-JUN-2002
DEFINITION	AX433201			
ACCESSION	AX433201			
VERSION	AX433201.1	GI:21658005		
KEYWORDS				
SOURCE	Bacillus licheniformis			
ORGANISM	Bacillus licheniformis			
REFERENCE	1			
AUTHORS	Berka, R. and Clausen, I. G.			
TITLE	Methods for monitoring multiple gene expression			
JOURNAL	Patent: WO 0229113-A 1616 11-APR-2002;			
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)				
FEATURES	Location/Qualifiers			
source	1..546			
	/organism="Bacillus licheniformis"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:1402"			
ORIGIN				
Query Match	18.5%;	Score 195.4;	DB 6;	Length 546;
Best Local Similarity	85.5%;	Pred. No. 2.6e-42;		
Matches	253;	Conservative	0;	Mismatches 37; Indels 6; Gaps 3;
Qy	766	CGCCAAATCGCTTTATTTATAAGAAAAACAGCAAGAAACAATCTATACATA---CCATT	822	
Db	27	CCAGAACCGCTTTATTTAAAGCAACCCGCCCAAGCAACATTTCTTATCCATCCATT	86	
Qy	823	ATGTAAGAGAGACACGAATCGCTGTACGCTTTTCGACAGGAG-AAAAATGCTG	881	
Db	87	AGSAAAGAGAGCGCCGAATGCTGTACGCTTTTCGCCAAGGAGAAAAAATGCGG	146	

```
Qy 882 TTTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCGTTCGCATCCTCGCT 941
|||
Db 147 TTTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCGTTCGCATCCTCGCT 206
|||
Qy 942 TCCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCT--CG 999
|||
Db 207 CCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTTTCGT 266
|||
Qy 1000 TCAATTTCCCGGTAGGCAAAAAACGCGAAGCGAGCATCATCTTGACCTGAAA 1055
|||
Db 267 CAAATTTNCCCGGTAGGCAAAAAACGCGAAGCGAGCATCATCTTGACCTGAAA 322
|||

RESULT 12
AE017012/c
LOCUS
DEFINITION Bacillus cereus ATCC 14579 section 15 of 18 of the complete genome.
ACCESSION AE017012 AE016877
VERSION AE017012.1 GI:29897926
KEYWORDS
SOURCE Bacillus cereus ATCC 14579
ORGANISM Bacillus cereus ATCC 14579
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
1 (bases 1 to 301332)
Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B.,
Kapatral,V., Bhattacharya,A., Reznik,G., Mikhailova,N.,
Lapidus,A., Chu,L., Mazur,M., Goltzman,E., Larsen,N., D'Souza,M.,
Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
Genome sequence of Bacillus cereus and comparative analysis with
Bacillus anthracis
JOURNAL Nature 423 (6935), 87-91 (2003)
PUBMED 12721630
REFERENCE
2 (bases 1 to 301332)
Candelon,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.
The number of ribosomal RNA operons in Bacillus cereus
Unpublished
JOURNAL
REFERENCE
3 (bases 1 to 301332)
Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B.,
Kapatral,V., Bhattacharya,A., Reznik,G., Mikhailova,N.,
Lapidus,A., Chu,L., Mazur,M., Goltzman,E., Larsen,N., D'Souza,M.,
Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
Direct Submission
JOURNAL Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de
Vilvert, Jouy en Josas 78352, France
FEATURES
source
1. 301332
/organism="Bacillus cereus ATCC 14579"
/mol_type="genomic DNA"
/strain="ATCC 14579"
/db_xref="ATCC:14579"
/db_xref="taxon:226900"
complement (114..470)
locus tag="BC4288"
complement (114..470)
locus tag="BC4288"
/codon_start=1
/transl_table=1
/product="Membrane-attached cytochrome c550"
/protein_id="AAP11201.1"
/db_xref="GI:29897927"
/transl_table="MKNRNLIPFALIALGIIIVMFVSFQGLNKSKEIADAKNGKPA
QTAGSKPDIYVQSCSTCHGDQLOAGVGNLQKIGKLSKDBEILSKGKNMPANIV
PADOAAKAVDWLSKKK"
complement (857..1984)
locus tag="BC4289"
complement (857..1984)
locus tag="BC4289"
/codon_start=1
/transl_table=1
/product="RNA polymerase sigma factor rpoD"

Qy 882 TTTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCGTTCGCATCCTCGCT 941
|||
Db 147 TTTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCGTTCGCATCCTCGCT 206
|||
Qy 942 TCCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCT--CG 999
|||
Db 207 CCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTTTCGT 266
|||
Qy 1000 TCAATTTCCCGGTAGGCAAAAAACGCGAAGCGAGCATCATCTTGACCTGAAA 1055
|||
Db 267 CAAATTTNCCCGGTAGGCAAAAAACGCGAAGCGAGCATCATCTTGACCTGAAA 322
|||

protein_id="AAP11202.1"
/db_xref="GI:29897928"
/translation="MADKPARSKQIETEMTLEQVKEQLTBLGKRGKGVLTVEEIAERNW
GFIESDQMDYEYIGEQIDLVGQNDNEEGNNRQISKTBEEDFNDLDSVPPGVKI
NDVRMYLKEIGRVDLLSABEEILATRIEEDDEAKRLAEANRLVSVIAKRYVGR
GMFLDLIQENMGLIKAVEKFDYRKGFESTVATWIRQAIIRAIQARTITIPVH
MVTINKLIRVQQLQDLGREPSEIPEIGMDLAPKEKVLAKIAOEPVSLTEPIGE
EDSHLGDFTIEDQATSPADHAAYELLKEOLEVDLDTLTDREENVLIRFGLDDGRTR
TLBEVGVFGVTRERINQIEAKALRKLHRSRSLKLDLFLE"
complement (2044..3840)
locus tag="BC4290"
complement (2044..3840)
locus tag="BC4290"
/EC number="2.7.7.-"
/codon_start=1
/transl_table=1
/product="DNA primase"
/protein_id="AAP11203.1"
/db_xref="GI:29897929"
/translation="MGNRIPEEVVQIQTSSDIVVEYVQLRKQGRNVFGLCPFHG
ENSPRSVSDKQIFHCFCGCGEGNVFSFLKMEGLAFTAVQKLGKRGIAVAEYTS
GQQQEDISDQTVIMQAHLLKKYHHLLVNTBEGNEALSILNLRGITKEMIKEFEI
GYASPAWDAATKILQKGLSLTMEQAGLLIRSEKDGSHYDFRGRVMPFTYIQGKV
IAPSGRALGDDTPKYNLSPETPIPHSKLLYNLPHQARPFIRKRGQVVLFGYADVLA
VKSQVEAVATMGALTTEQAKLLRNVTVLVLCYDGDKAGREATKAGOLLKLVGCO
VKYTSLPDKLDPDEYVQOYCTAFENLVKSVSVFVGKYNLYRLGNLQDSSGKEEV
KSVLEKSLQLQMAQESYKLSQESFSYMETLLNLQHQYRKQKQKQKQKQKQKQK
SQVQTKPLKTGFERAEREIYYHMLQSAVAVRMESHEDFPHTEHKGILLYELVAYE
KGNPSVGTFLSWSLDEKLNIIITDIDTDEFINPEYTEVQLQGHLETLRHQRLEKM
EIIPIKVKQMEKTDPEAAKYVAYLQNKARK"
4007..4213
locus tag="BC4291"
4007..4213
locus tag="BC4291"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP11204.1"
/db_xref="GI:29897930"
/translation="MLYKIVYICSSVTNLSRNNNTYTPFYKSGSKPHILFDLEICNFF
MLHITINSILFSYKDPALILFP"
complement (4202..4333)
locus tag="BC4292"
complement (4202..4333)
locus tag="BC4292"
/codon_start=1
/product="EMG2 protein"
/protein_id="AAP11205.1"
/db_xref="GI:29897931"
/translation="MQNINILVGVDPYVYKAKIVQVGAEPHVGLFVASYAHLPRK"
complement (4701..5513)
locus tag="BC4293"
complement (4701..5513)
locus tag="BC4293"
/codon_start=1
/transl_table=1
/product="hypothetical Cytosolic Protein"
/protein_id="AAP11206.1"
/db_xref="GI:29897932"
/translation="MDNKKIVVYVSDSVGETADLVVRAAMQGFPEAPDIRRVYVEDTGT
TLKEVSIKSNQALICFTLVKQDMROYLATEAKGVEAYDIIIGPLIDQIEETIGQV
PRYEPGVRLDEYEFYKIEAIEFPAVKYDGRDARGILKADIVLIGVSRSTKTLPSQY
LAHNKRVANVPLVPEVDPPELYQVAKGFCGLKITPDKLNHIRKERLSKSLGSDG
ATYANIRIGEEIDHFEVIRINQVIDVSNKAIBETANIIVNAVQNKQMF"
complement (5543..6175)
locus tag="BC4294"
complement (5543..6175)
locus tag="BC4294"
/codon_start=1
/transl_table=1
/product="CBS domain containing protein"
/protein_id="AAP11207.1"
```

```
/db_xref="GI:29897933"
/translation="MIITELNKRQEHIIQIVKDHGPIITGESIAAOLGLTRATLPDLA
ILTMAGYLEARPRVGYFTGKGQLLSEAVKKIKVQDYQSRPVVIDKNVSYDAICT
MFLBVDGLTVVDQMTLLVGVSRKDLRLASLGKDLTSLPWNILMTMPNIMACRRE
DSLYDIAMELIERQIDAMPVKDKTQGLEVIGRIITKNTITRAFVNLVNE"
complement(6283..7029)
/locus_tag="BC4295"
complement(6283..7029)
/locus_tag="BC4295"
/codon_start=1
/transl_table=11
/product="DNA repair protein recO"
/protein_id="AAP11208.1"
/db_xref="GI:29897934"
/translation="MFQKVEGIVRTTDDYGETNKNIVTIFSRBLKVSAMARGSKPKS
RLASVQLMTHGHFIQMGSLGTQQGHEIISTMKEIREIDPLTAYASFIVELTDKAT
EKKHNPYLFEMLYOTLHYMCEGVDPVLSLIYQTKMLPVLGMRPVFTFCACHOETD
FVAFSVRGGFLCSRHSQDPYRIIPVGEAVHKLKLLFFHFDLHRLGNVSRKSTKKQM
RLVMTYDVEYCGIYKSRFLFELQDKFQI"
complement(7063..7206)
/locus_tag="BC4296"
complement(7063..7206)
/locus_tag="BC4296"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP11209.1"
/db_xref="GI:29897935"
/translation="MLDTWKFFSGTSGIETVLLKMEKQVNDQHEBELAHLDS
FIS"
complement(7383..8288)
/locus_tag="BC4297"
complement(7383..8288)
/locus_tag="BC4297"
/codon_start=1
/transl_table=11
/product="GTP-binding protein"
/protein_id="AAP11210.1"
/db_xref="GI:29897936"
/translation="MNRKYGKSGVSIIGRPNVGKSTFLNRIIGOKIAMSCKPOTTR
NKIQGVYENDAQVIFIDTGHGKPKHLGDPVKAQOTLKEVDIVLPMVNAVGEFG
RGEFTIEKLETKQPVFLVINKIDLHPEQLLELDQYRKLHFAEIVPISALDGN
VDALGTIKYLPFGPYPPENQVTDHPERFIIASLIREKVLHLTRSVPHSVAVVID
AIQKREGAVINATIVVERASQGIIGKQGMKLEKVGKARFDIEALLGSKVLEY
WVKQKWRNKMSQLRDIGFREDEY"
complement(8281..8679)
/locus_tag="BC4298"
complement(8281..8679)
/locus_tag="BC4298"
/EC_number="3.5.4.5"
/codon_start=1
/transl_table=11
/product="Cytidine deaminase"
/protein_id="AAP11211.1"
/db_xref="GI:29897937"
/translation="MNSQLILEAEARQAQYVPYKFGVGAALITRGKVVYRGCNVE
NASYGLNCNAERTALFKAVSEGDFEVAIVVADTKRPVPCGACRQVWVELCKQDTK
VYLSLHGDVQSTTVGELLPGAFLAEDLHE"
complement(8780..9133)
/locus_tag="BC4299"
complement(8780..9133)
/locus_tag="BC4299"
/EC_number="2.7.1.107"
/codon_start=1
/transl_table=11
/product="Diacylglycerol kinase"
/protein_id="AAP11212.1"
/db_xref="GI:29897938"
/translation="MKKGKLIIDFGVAIAGIFFCFRHERNMKIHVLAIVVICGFFYF
15.7%; Score 166; DB 1; Length 301332;
Best Local Similarity 62.8%; Pred. No. 5e-34;
Matches 275; Conservative 0; Mismatches 160; Indels 3; Gaps 1;
```

```
QY 621 TTCTAGCTTTGCCTGTTTGTGTTTCAATGAAGACAGCTCAAGAGGTGAATCAACGTTG 680
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124209 TTCTATCATTAATATGTTTAAATGGAATACATGATTTTAAGAGGAGAGATCGTATTTTG 124150
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 681 ATCGAATTCGTAAGAGGACGATTGATTTATGTATCGCCCAATATATTTGTCATTTGAAAC 740
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124149 TTTGAATATGTTTACAGGTTTACGTGGAGTATGTAGGGCCGGAATATGTCGTAATTGATCAT 124090
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 741 GCGGGGATCGGCTATCAGATCTTTCACGCCAAATCCGTTTATTATTAAAGAAAAACAGCAA 800
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124089 AATGGAATTTGGTTATCAAAATTTTACACCAAAATCCGTATGTTATTTCAAGAAGATGACAA 124030
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 801 GAAA---CAATCTATACATACCATTTATGTAAGAGAAGACACGAATTCGCTGTACGGCTTT 857
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124029 GAAATCCGTGCTATACATATCATTTATGTGAGAGAAGATATTTATGGCACTTTTACGGGTTT 123970
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 858 TCGCAAGGGGAAGAAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGGATCGGCCCA 917
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123969 AAAACACGTGAAGAGCGTTTATTATTATTAAGATTTGTTAGGCGTGTCTGGCAATGGGCCA 123910
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 918 AAAGGAGCGTTTCGATCCTCGCTTCGCGCATCCGGAGCGGTGATTGAACGATCGAG 977
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123909 AAAGGAGCTCTTGCATTTTACGTTTCTGTTCAAACAGGACAGGTGTTCAACCGATTGAG 123850
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 978 AATGAGGACGAAGCAATTTCTCGTCAAAATTTCCCGCGGTAGGCACAAAAACGGCAAGGCCAG 1037
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123849 CATGAAGATGAGAAGTTTATTAGTGAATTTCCAGGCGTCCGAAAGAAAACACACGCCCA 123790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1038 ATCATCTTGACCTGCTGAAA 1055
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123789 ATGATTTTAGATTTAAAA 123772
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
AE017278/c
LOCUS
DEFINITION
  AE017278
  Bacillus cereus ATCC 10987, section 15 of 18 of the complete
  genome.
ACCESSION
  AE017278
  AE017194
VERSION
  AE017278.1
  GI:42739327
KEYWORDS
  Bacillus cereus ATCC 10987
SOURCE
  Bacillus cereus ATCC 10987
  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
  cereus group.
REFERENCE
  1 (bases 1 to 288814)
  Rasko,D.A., Ravel,J., Okstad,O.A., Helgason,E., Cer,R.Z., Jiang,L.,
  Shores,K.A., Fouts,D.E., Tourasse,N.J., Angiuoli,S.V., Kolonay,J.,
  Nelson,W.C., Kolsto,A.-B., Fraser,C.M. and Read,T.D.
  The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
  adaptations and a large plasmid related to Bacillus anthracis pX01
  Nucleic Acids Res. 32 (3), 977-988 (2004)
JOURNAL
  PUBLISHED
  14960714
REFERENCE
  2 (bases 1 to 288814)
  Rasko,D.A., Ravel,J., Okstad,O.A., Helgason,E., Cer,R.Z., Jiang,L.,
  Shores,K.A., Fouts,D.E., Tourasse,N.J., Angiuoli,S.V., Kolonay,J.,
  Nelson,W.C., Kolsto,A.-B., Fraser,C.M. and Read,T.D.
  Direct Submission
  Submitted (19-FEB-2004) The Institute for Genomic Research, 9712
  Medical Center Dr, Rockville, MD 20850, USA
FEATURES
  Location/Qualifiers
  1..288814
  /organism="Bacillus cereus ATCC 10987"
  /mol_type="genomic DNA"
  /strain="ATCC 10987"
  /db_xref="ATCC:10987"
  /db_xref="taxon:22523"
  complement(192..2312)
  /locus_tag="BCE4353"
  complement(192..2312)
  /locus_tag="BCE4353"
  /note="Identified by match to protein family HMM PF00905;
```

match to protein family HMM PF03717"  
/codon\_start=1  
/transl\_table=11  
/product="penicillin-binding protein"  
/protein\_id="AAS43254.1"  
/db\_xref="GI:42739328"  
/translation="MGKKQKQKKTHVPFRLNLVFCVLMFSAVIRLGYQIVRG  
EYKNEVERKENSTISNPVRGKIFDRYGRPVVDNAARTITFTKMGSTAEARELETA  
KKLADLLEVPDLDLDRDKDYWLAIHKEEAQEKITKDRQEFKDKIDDKELDERQR  
NRVTEEVNQSLAKDEVLAIKSMQGYAMTPQVIKKDATEQYALISENLAALPGV  
DTNVDORKYYDDLFRSLVSGVSTDEGLPRERLDYLVDRYNRNDRVGRKSYLEQQY  
EDSHGKAERVNITDKNGNLETINVSKGSGNNLALTIDMELOKRVETITKNLLR  
YKGGPILLDRVFMVMPKNGEVLMSMAGLQVEENGKTKQDFALGTMTWTSYPMGSTV  
KGATLVGYQTKAQPSYQLDDEPIVLKGTPEKSSMTGMSINDLTALKSSNVYMEK  
TAMNIAQVQYRGCTLDIPQKAPDTMRYFFQCGGLGVKGTGIDLLENESAGQTGKGIQPG  
FLDLSTGYDTYTPQLAQKVTSTANGYRQPVQVKIRQPEAKQGEVGVKVSME  
PKVLRNIMPESQIKRQEGFRQFVNDSSGATPKYFAGAPYKAAAGKTGTATQVYGGDK  
EIGRKANGERKETYNLTLVGYAPLEDPEVAFSVVVPVVDKSGINGVISRIMDAYFD  
LKKVGENGEATQEEIDKKNKQDDE"  
complement (2422..3699)  
/locus\_tag="BCE4354"  
complement (2422..3699)  
/locus\_tag="BCE4354"  
/codon\_start=1  
/transl\_table=11  
/product="membrane protein, putative"  
/protein\_id="AAS43255.1"  
/db\_xref="GI:42739329"  
/translation="MKWKHHIDGVEVNRDLVLLLIGLYTLAISLNTFVNIYLMKQ  
TQNTVNLGNLASVYVQLPTFLIGGLAKRIDRSILIRIGVGTLAIFVIVLVAGKS  
ASHYLLGALLGVGYWLAFLNLTPEITPETRDFPFNGFLGLTSPSGMIGPIAA  
GYTIRMEKSGSYTVIRPFLSLLPAIAVLVSFFYKRECEGRYEVAOVLKERRIDKMK  
GRITRAPHFQGLRGCTFIPVSVYVYLATDSELALGXSLVNSAVSFVYVLVARMK  
KEWRKALLGGLIUYAVFLVIFNVYTKLLIYAACIAIAYIPILLVPYGSMTYDVIG  
RAKNAERVRVYVVRVRELNAGRICSVLSLCAVLPFPPEKSLPFLLCILAGHFLI  
YFAVNVKRYDEGNAGNTSVVAQGTQNTQTEPEG"  
complement (3816..4427)  
/gene="sodA"  
/locus\_tag="BCE4355"  
complement (3816..4427)  
/gene="sodA"  
/locus\_tag="BCE4355"  
/EC\_number="1.15.1.1"  
/note="identified by match to protein family HMM PF02777"  
match to protein family HMM PF02777"  
/codon\_start=1  
/transl\_table=11  
/product="superoxide dismutase, Mn"  
/protein\_id="AAS43256.1"  
/db\_xref="GI:42739330"  
/translation="MAKHENLPYAYDALEPHDPDKETMNIHHTKHNTYITNLNAAL  
EGHAELADKSVLEELVNLNEVPEAIRAVRNGGGHANHTFFWTILSPNGGGQPVGEL  
ATAIEAKFGSPDAFKEEFKAGATRFSGSMAWLVNNGELEVTSTPNQDSLTEGKTP  
VIGLDVWEHAYLNYQNRRPDYIGAFWNVDWNAEKRYQAK"  
4649..4722  
/gene="tRNA-Met"  
/locus\_tag="BCE5708"  
4649..4722  
/gene="tRNA-Met"  
/locus\_tag="BCE5708"  
/product="tRNA-Met"  
/product="tRNA-Met"  
4965..5729  
/locus\_tag="BCE4356"  
4965..5729  
/locus\_tag="BCE4356"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAS43257.1"  
/db\_xref="GI:42739331"  
/translation="MSIFQLVKSVPKDMALFRFQKIGKTIYILMLCLITPIPT  
FFYGSFIQDSVTMVAIEKDLPDFKIENGELKADIEQPIQKEGDLAFVDPDNTDI  
EKYQNKTLGLFKDKVWSIGNGQTQYSYNDLLGTSLEKDLQDFISVFDISIYPIILA

VIGLVVLFOLFITFLGITLLAFIGSAMSGORKLSYKQVMTLTAYSYTIPITFFMIMD  
LFKIVPGSGTFIYIAVVLIVLTIKEVPRNEKNA"  
5783..6148  
/locus\_tag="BCE4357"  
5783..6148  
/locus\_tag="BCE4357"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAS43258.1"  
/db\_xref="GI:42739332"  
/translation="MKELGFLFLVLVILGYIFYFYDIKIGTLPMLSSYKKTNAETIKQE  
STNKOKENKAETDVTYKTEVKTGDTVLSITEANKKPSIEKVIDDFKQLNK  
NTPATKIQLGSKSYKFPYQ"  
6313..7416  
/gene="gcpB"  
6313..7416  
/locus\_tag="BCE4358"  
6313..7416  
/gene="gcpB"  
/locus\_tag="BCE4358"  
/note="identified by match to protein family HMM PF01077;  
match to protein family HMM PF04551; match to protein  
family HMM TIGR00612"  
/codon\_start=1  
/transl\_table=11  
/product="gcpB protein"  
/protein\_id="AAS43259.1"  
/db\_xref="GI:42739333"  
/translation="MTHRTKRPVKVGNLTIGGNELIIOQSMTTTKTHDVBATVAEIK  
RLBAGQVVRVAVPDRAANAADIKKQINIPVADHFDYRLAKAIEGGIDKVR  
IPNGNIGRRHVAVVNAKRGIPIRIGVNAKSERHLEKYGYPADGMVESALHHI  
KILEDLDFHDIIVSMKASDVNLAEAVEKARAFDYPHLGITEGTLFACTGVKSAAG  
LGAILSGINTLRISLSADPVEEVKVARRELLKSPGLASNAATLISCTCGRIDILL  
SIANEVBEISTIQVPKVAVLGCAVNGPGAREADIGIARGEGGLFRKQVVRKV  
PESIMVBEELKXEDIVTAAEAAREKEKETQE"  
complement (7502..7915)  
/locus\_tag="BCE4359"  
complement (7502..7915)  
/locus\_tag="BCE4359"  
/note="identified by match to protein family HMM PF01475"  
/codon\_start=1  
/transl\_table=11  
/product="transcriptional regulator, Fur family"  
/protein\_id="AAS43260.1"  
/db\_xref="GI:42739334"  
/translation="MNLTEALRLMKDKYKHTGKREMLRFLFAHNRYLTAKDVLEHM  
KDDYPGLSFDITVYRNLTVPABIGVLEQTELNGEKHFRTCSIMEHHHHFICLDCGGTK  
EITSCPMDFMKDFGVEYTGKFEIYGRCPKCAK"  
complement (7929..8762)  
/locus\_tag="BCE4360"  
complement (7929..8762)  
/locus\_tag="BCE4360"  
/note="identified by match to protein family HMM PF00950"  
/codon\_start=1  
/transl\_table=11  
/product="cation ABC transporter, permease protein,  
putative"  
/protein\_id="AAS43261.1"  
/db\_xref="GI:42739335"  
/translation="MIQDFLQYDFLRNSLYAGILIGLVAPLIGVFWVIRRMSLIADAL  
SHVTLSGIAASLLLEKTFIQGFLNPLYMGMI FSGGALLIEKRLTVKHYQELAIPI  
ILSAGMGIVIPITSLANGENTDLPSYLPFGSVSAVSTDLIIIGIVAIWIVITLLYK  
ELFLLSPDEEYANSTGLSAKNIHFIPIILVALVAVSMRVVGVLLVSLMTLPVAASI  
RIANGFKQTIFFSILFGEIYVIGGFASYQLDLAPGGTIVMIAVLILGAILNWKKT  
A"  
complement (8762..9532)  
/locus\_tag="BCE4361"  
complement (8762..9532)  
/locus\_tag="BCE4361"  
/note="identified by match to protein family HMM PF00005"  
/codon\_start=1  
/transl\_table=11  
/product="cation ABC transporter, ATP-binding protein,



putative"  
/protein\_id="AAS43262.1"  
/db\_xref="GI:42739336"  
/translation="MNNILIEGLSFYEDRNVLIEDINLQVPKGAFLGLVGPNGSGKS  
TLKCLGLVLPKQSIKILFGVDSKKFKEMKNKGVYSOKANSFNSGRPATVFEVVSMSG  
LVSKGLFRFFTKNDKKVKRAIDVGNSEFGNIGELSGQQORVFIAIALVSDPE  
LLILDEPTVGDIVKRVSPFYLEILDKRLGITLILVTHDMGAVTEKVTHTVACLNOHL  
HFHGNVEXFRELEDAEMSVLYGHVHRLHEHEHHGRI"  
complement(9713..10591)  
/locus\_tag="BCE4362"  
complement(9713..10591)  
/locus\_tag="BCE4362"  
/note="identified by match to protein family HMM PF02588"  
/codon\_start=1  
/transl\_table=11  
/product="membrane protein, putative"  
/protein\_id="AAS43263.1"

Query Match 15.6%; Score 164.4; DB 1; Length 288814;  
Best Local Similarity 62.6%; Pred. No. 1.4e-33;  
Matches 274; Conservative 0; Mismatches 161; Indels 3; Gaps 1;  
QY 621 TTCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGTCAAGAGGTGAATGAACGTTG 680  
Db 138093 TTCTATCATTAATATGTTAAATGGAATACATGATTTAAGAGGAGAGATTTGTTTG 138034  
QY 681 ATCGAAATCGTAAAGGACGATGATGATGATCGCCCAATATATTGTCATTGAAAC 740  
Db 138033 TTTGAATATGTTACAGGTTACGTGAGTATGTAGGACCGGAATATGTCGTAATGATCAT 137974  
QY 741 GCGGGATCGGCTATCAGATCTTACGCGCAATCGGTTTATTATTAAAGAAACAGCAA 800  
Db 137973 AATGGAATTGGTATPACAAATTTTACACCGCAATCGGTATGATTTTCAAGAGTAAGCAA 137914  
QY 801 GAAA---CAATCTATACATACATGATGATGAAGAGACACGAATGCGCTGTACGCTTT 857  
Db 137913 GAATCCGCTGCTATACATATCAATATGATGAGAGAGATATATTGCGGCTTTACGGGTTT 137854  
QY 858 TCGAAGAGGAGAAATAATGCTGTTTACGAAATGCTGAATGTTTACGGGGATCGGCCA 917  
Db 137853 AAAACACGTGAAGACGTTTATTATTACAAAGTTGTTAGTGCTGTTGATTTGGACCA 137794  
QY 918 AAAGAGCGCTTGCATCCTGCTTCCGGCGATCCGGGAGCGGTGATTTGAAGCGATCGAG 977  
Db 137793 AAAGCGCTTTGCAATTTTAGCTTCTGCTCAAAATTTCCGCGGTAGGCMAAAACCGCAAGGCG 137734  
QY 978 AATGAGGACGAGCATTTCTGCTCAATTTCCGCGGTAGGCMAAAACCGCAAGGCGAG 1037  
Db 137733 CATGAAGACGAGAAAGTTTATTAGTGAAGTTCCGCGCGCTCGGAAAGAAACACGACGCCAA 137674  
QY 1038 ATCATCCTTGACCTGAAA 1055  
Db 137673 ATGATTTTAGATTTAAAA 137656

RESULT 14

AE017225\_42/c  
WPCOMMENT

Sequence split into 53 fragments LOCUS AE017225 Accession AE017225

Fragment Name	Begin	End
AE017225_00	1	110000
AE017225_01	100001	210000
AE017225_02	200001	310000
AE017225_03	300001	410000
AE017225_04	400001	510000
AE017225_05	500001	610000
AE017225_06	600001	710000
AE017225_07	700001	810000
AE017225_08	800001	910000
AE017225_09	900001	1010000
AE017225_10	1000001	1110000
AE017225_11	1100001	1210000
AE017225_12	1200001	1310000

1300001 1410000  
AE017225\_13 1510000  
AE017225\_14 1610000  
AE017225\_15 1710000  
AE017225\_16 1810000  
AE017225\_17 1900001  
AE017225\_18 2000001  
AE017225\_19 2100000  
AE017225\_20 2200001  
AE017225\_21 2310000  
AE017225\_22 2400001  
AE017225\_23 2500001  
AE017225\_24 2600001  
AE017225\_25 2710000  
AE017225\_26 2800001  
AE017225\_27 2900001  
AE017225\_28 3000001  
AE017225\_29 3110000  
AE017225\_30 3210000  
AE017225\_31 3300001  
AE017225\_32 3400001  
AE017225\_33 3500001  
AE017225\_34 3600001  
AE017225\_35 3710000  
AE017225\_36 3810000  
AE017225\_37 3900001  
AE017225\_38 4000001  
AE017225\_39 4100000  
AE017225\_40 4200001  
AE017225\_41 4310000  
AE017225\_42 4400001  
AE017225\_43 4500001  
AE017225\_44 4600001  
AE017225\_45 4700000  
AE017225\_46 4800001  
AE017225\_47 4900001  
AE017225\_48 5000001  
AE017225\_49 5100000  
AE017225\_50 5200001  
AE017225\_51 528663  
AE017225\_52  
Continuation (43 of 53) of AE017225 from base 4200001 (AE017225 Bacillus anthracis str.

Query Match 15.3%; Score 161.2; DB 1; Length 110000;  
Best Local Similarity 62.1%; Pred. No. 9.7e-33;  
Matches 272; Conservative 0; Mismatches 163; Indels 3; Gaps 1;  
QY 521 TTCTAGCTTTCCCTGTTTGTGTTTACATGAAGAGCAGTCAAGAGGTGAATGAACGTTG 680  
Db 28998 TTCTATCATTAATATGTTAAATGGAATACATGATTTAAGAGGGAGAGATTTGTTTGT 28839  
QY 681 ATCGAAATCGTAAAGGACGATGATGATGATCGCCCAATATATTGTCATTGAAAC 740  
Db 28938 TTTGATATGTTACAGGTACGTGAGTATGAGACCGGAATATGCTGTAATGATCAT 28779  
QY 741 GCGGGATCGGCTATCAGATCTTACGCGCAATCCGTTTATTATTATTAAGAAACAGCAA 800  
Db 28778 AATGGAATGCGTATCAAAATTTTCAACCGAATCGGTATGATTTTCAAGAGTAAGCAA 28719  
QY 801 GAAA---CAATCTATACATACATGATGATGAAGAGACACGAATGCGCTGTACGCTTT 857  
Db 28718 GAAATCCGCTGCTATACATATCATATGATGAGAGAGATATATTGCGCACTTTACGGGTTT 28659  
QY 858 TCGAAGAGGAGAAATAATGCTGTTTACGAAATGCTGAATGTTTACGGGGATCGGCCA 917  
Db 28658 AAAACACGTGAAGACGCTTTATTATTATTACAAGTTGTTAGGGGTGCTGTTGTTGACCA 28599  
QY 918 AAAGAGCGCTTGCATCCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAG 977  
Db 28598 AAAGGTGCTCTTGCAATTTTAGCTTCTGGTCAACAGGACAGGTGCTTCAAGCGATTGAA 28539  
QY 978 AATGAGGACGAGCATTTCTGCTCAAAATTTCCGCGCGTAGGCMAAAACCGCAAGGCGAG 1037  
Db 28538 CATGAAGACGAGAAAGTTTATTAGTGAAGTTCCGCGCGCTCGGAAAGAAACACGACGCCAA 28479

Qy 1038 ATCATCCTTGACCTGAAA 1055  
Db 28478 ATGATTTTAGATTTAAAA 28461

RESULT 15  
AE017334\_42/c  
WPCOMMENT

Sequence split into 53 fragments LOCUS AE017334 Accession AE017334

Fragment Name	Begin	End
AE017334_00	1	110000
AE017334_01	100001	210000
AE017334_02	200001	310000
AE017334_03	300001	410000
AE017334_04	400001	510000
AE017334_05	500001	610000
AE017334_06	600001	710000
AE017334_07	700001	810000
AE017334_08	800001	910000
AE017334_09	900001	1010000
AE017334_10	1000001	1110000
AE017334_11	1100001	1210000
AE017334_12	1200001	1310000
AE017334_13	1300001	1410000
AE017334_14	1400001	1510000
AE017334_15	1500001	1610000
AE017334_16	1600001	1710000
AE017334_17	1700001	1810000
AE017334_18	1800001	1910000
AE017334_19	1900001	2010000
AE017334_20	2000001	2110000
AE017334_21	2100001	2210000
AE017334_22	2200001	2310000
AE017334_23	2300001	2410000
AE017334_24	2400001	2510000
AE017334_25	2500001	2610000
AE017334_26	2600001	2710000
AE017334_27	2700001	2810000
AE017334_28	2800001	2910000
AE017334_29	2900001	3010000
AE017334_30	3000001	3110000
AE017334_31	3100001	3210000
AE017334_32	3200001	3310000
AE017334_33	3300001	3410000
AE017334_34	3400001	3510000
AE017334_35	3500001	3610000
AE017334_36	3600001	3710000
AE017334_37	3700001	3810000
AE017334_38	3800001	3910000
AE017334_39	3900001	4010000
AE017334_40	4000001	4110000
AE017334_41	4100001	4210000
AE017334_42	4200001	4310000
AE017334_43	4300001	4410000
AE017334_44	4400001	4510000
AE017334_45	4500001	4610000
AE017334_46	4600001	4710000
AE017334_47	4700001	4810000
AE017334_48	4800001	4910000
AE017334_49	4900001	5010000
AE017334_50	5000001	5110000
AE017334_51	5100001	5210000
AE017334_52	5200001	5227419

Continuation (43 of 53) of AE017334 from base 4200001 (AE017334 Bacillus anthracis str.

Query Match 15.3%; Score 161.2; DB 1; Length 110000;  
Best Local Similarity 62.1%; Pred. No. 9.7e-33;  
Matches 272; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

Qy 621 TTCTAGCTTTTGCTCTTTTGTGTACAATGAAGAGCAGTCAAGAGGTGAATGAACGTTG 680  
Db 28526 TTCTATCATTAATATGTTTAAATGAATACATGATTTAAGAGGGAGAGATTGTATTTTG 28467

Qy 581 ATCGAATTCGTAAGGACGATTGATTATGATCGCCCAATATATGTCATTGAAAAC 740  
Db 28466 TTTGAATATGTTACAGGTTACGTGGAGTAGTAGGCCGGAATATGCGTAATTGATCAT 28407  
Qy 741 GCGGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTATAAGAAAAACAGCAAA 800  
Db 28406 AATGGAATTGGTTATCAAATTTTTCACACCGNATCCGTATGTATTTCAAAGAAGTAAGCAA 28347  
Qy 801 GAAA---CAATCTATACATACCATTTATGTAAGAGAAACACACGAATGCGCTGTACGGCTTT 857  
Db 28346 GAAATCCGTCTCTATACATATCATTTATGTGAGAGAGATATTATGGCACTTTTACGGGTTT 28287  
Qy 858 TCGACAGGAGNAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCCA 917  
Db 28286 AAACACGTGAAGAGCGTTTATTATTACAAAGTTGTAGGGGTGCTGGTATTGGACCA 28227  
Qy 918 AAAGGAGCGCTTGCGATCCTCGCTTCGGGCGATCCGGGAGCGGTGATTGAAGCGATCGAG 977  
Db 28226 AAGGTGCTCTTGCAATTTTAGCTTCTGGTCAACAGGACAGGTCGTTCAAGCGATTGAA 28167  
Qy 978 AATGAGACGAAAGCATTTTCTCGTCAAAATTTCCCGGCGTAGGCAAAAAACCGCAAGGAG 1037  
Db 28166 CATGAGACGAGAAAGTTTATTAGTGAAGTTCGCGGCGCTCGGAAAGAAAAACAGCAGCCAA 28107  
Qy 1038 ATCATCCTTGACCTGAAA 1055  
Db 28106 ATGATTTTAGATTTAAAA 28089

Search completed: February 27, 2006, 08:07:00  
Job time : 5721 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:32:21 ; Search time 4975 Seconds  
(without alignments)  
9921.683 Million cell updates/sec

Title: US-10-510-408-1

Perfect score: 1055

Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	6.6	767	10	CL678407
2	68.2	6.5	524	2	BG455392
3	64	6.1	790	9	BH371812
4	49.2	4.7	707	8	CX137869
5	47	4.5	468	8	CX067472
6	47	4.5	485	8	CX067473
7	45.8	4.3	745	7	CX384972
8	44.6	4.2	1101	10	CNS0039G
9	42.8	4.1	501	5	BQ496577
10	42.8	4.1	519	5	BQ501436
11	42.8	4.1	527	6	CB853944
12	42.4	4.0	684	8	CX146914
13	42.4	4.0	701	8	CX144590
14	42.4	4.0	738	8	CX153248
15	41.2	3.9	882	9	AZ674153
16	41	3.9	859	9	AZ540067
17	41	3.9	919	9	BH139275
18	40.8	3.9	451	10	AG936518
19	40.8	3.9	468	10	AG938690
20	40.8	3.9	493	10	AG911866
21	40.8	3.9	546	10	AG944693
22	40.8	3.9	566	10	AG939340

C 23	40.8	3.9	585	10	AG910157
C 24	40.8	3.9	672	10	AG920520
C 25	40.4	3.8	1101	10	CNS0100X
C 26	40	3.8	470	1	AJ761112
C 27	39.8	3.8	405	3	BP813221
C 28	39.8	3.8	796	8	CX949675
C 29	39.8	3.8	855	8	DN640592
C 30	39.8	3.8	905	8	DR887954
C 31	39.6	3.8	338	5	BW578460
C 32	39.6	3.8	542	5	BW585582
C 33	39.4	3.7	374	9	BZ357018
C 34	39.4	3.7	828	7	CK090955
C 35	39.2	3.7	386	10	CZ673753
C 36	39.2	3.7	395	3	BP798691
C 37	39.2	3.7	400	2	BE521520
C 38	39.2	3.7	519	3	BP562728
C 39	39.2	3.7	587	7	CO221100
C 40	39.2	3.7	604	8	DR540570
C 41	39.2	3.7	642	3	BP796746
C 42	39.2	3.7	1524	4	CNS0ADNT
C 43	39.2	3.7	1713	4	CNS0ADK8
C 44	38.8	3.7	1101	10	CNS0125V
C 45	38.8	3.7	1350	10	CL019486

#### ALIGNMENTS

##### RESULT 1

##### CL678407

##### LOCUS

##### DEFINITION

CL678407 767 bp DNA linear GSS 09-JUL-2004  
PRI0122d\_P08.2 - PRI0122d\_BR (767) Mixed stage fosmid library of P.  
pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.

##### ACCESSION

##### VERSION

##### KEYWORDS

##### SOURCE

##### ORGANISM

##### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### PUBMED

##### COMMENT

CL678407 767 bp DNA linear GSS 09-JUL-2004  
PRI0122d\_P08.2 - PRI0122d\_BR (767) Mixed stage fosmid library of P.  
pacificus var. California Pristionchus pacificus genomic, genomic  
survey sequence.  
CL678407 GI:50184707  
GSS.  
Pristionchus pacificus  
Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.  
1 (bases 1 to 767)  
Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.  
AppADB: an AcedB database for the nematode satellite organism  
Pristionchus pacificus  
Nucleic Acids Res. 32 (1), D421-D422 (2004)  
14581447  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
This library was generated at Caltech, Pasadena, USA and end  
sequenced at Vancouver, Canada.  
Seg primer: T7  
Class: fosmid ends.  
Location/Qualifiers  
1. 767  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Mixed stage fosmid library of P. pacificus  
var. California"  
/note="Vector: pEpifos-5 Fosmid vector"

##### ORIGIN

Query Match 6.6%; Score 70; DB 10; Length 767;  
Best Local Similarity 54.5%; Pred. No. 2.3e-08;  
Matches 140; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

799 AAGAAACAACTATACATACCATATATAGAGAGAGACAGATGCCTGTACGGCTTTT 858  
 506 AAGCGATCGTTTTCACCCACTTTTGTGGTGAAGACGGCAGCTGCTGTACGGTTTAA 565  
 859 CGACAAGGAGAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCCAA 918  
 566 ACAATAAACAGAGCGCAGTGTTCAGAGAGTTGATCAAAACACGCGCTCGGCCCGA 625  
 919 AAGGAGCGCTTGGCATCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCAGA 978  
 626 AGTTGGCGCTGGCGATCTCTCCGGAATGTACGCGCAGCACTTGTGTAATGCCGTGAGC 685  
 979 ATGAGGAGCAAGCAATTTCTGTCATAATTTCCGGCGTAGGCAAAAAACGCGCAGGCGA 1038  
 686 GTGAAGAGTGGGGCGCTGTGTAACTGCGGGTATTGGCAAAAAAACCGCGCAACGCT 745  
 1039 TCATCCTTGACCTGAAA 1055  
 746 TGATTTGTAATGAAA 762

RESULT 2  
 BG455392/c  
 LOCUS  
 DEFINITION  
 BG455392 524 bp mRNA linear EST 19-MAR-2001  
 clone NF047E06PL 5', mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Medicago truncatula (barrel medic)  
 Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 REFERENCE  
 AUTHORS  
 Liu J., Scott A.D., Harris A.R., Gonzales R.A., Bell C.J.,  
 Flores H.R., Iman J.T., Weller J.W., May G.D. and Harrison M.J.  
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 Medicago truncatula phosphate-starved leaf library  
 Unpublished (2000)  
 CONTACT: Harrison MJ  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7325  
 Fax: 580 221 7380  
 Email: mjharrison@noble.org  
 Insert Length: 524 Std Error: 0.00  
 Plate: 047 row: E column: 06  
 Seq primer: TCACACAGGAACAGCTATGAC.  
 Location/Qualifiers  
 1..524  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF047E06PL"  
 /tissue\_type="leaf"  
 /dev\_stage="trifoliolate"  
 /clone\_lib="Phosphate starved leaf"  
 /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20mM potassium phosphate. RNA was prepared from above ground tissues."

FEATURES  
 source  
 1..524  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3880"  
 /clone="NF047E06PL"  
 /tissue\_type="leaf"  
 /dev\_stage="trifoliolate"  
 /clone\_lib="Phosphate starved leaf"  
 /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20mM potassium phosphate. RNA was prepared from above ground tissues."  
 ORIGIN  
 Query Match 6.5%; Score 68.2; DB 2; Length 524;  
 Best Local Similarity 54.1%; Pred. No. 6.8e-08;  
 Matches 139; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
 799 AAGAAACAACTATACATACCATATATAGAGAGAGACAGATGCCTGTACGGCTTTT 858

438 AAGCGATCGTTTTCACCCACTTTTGTGGTGAAGACGGCAATGCTGTACGGTTTAA 379  
 859 CGACAAGGAGAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCCAA 918  
 378 ACAATAAACAGAGCGCAGTGTTCAGAGAGTTGATCAAAACACGCGCTCGGCCCGA 319  
 919 AAGGAGCGCTTGGCATCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCAGA 978  
 318 AGTTGGCGCTGGCGATCTCTCCGGAATGTACGCGCAGCACTTGTGTAATGCCGTGAGC 259  
 979 ATGAGGAGCAAGCAATTTCTGTCATAATTTCCGGCGTAGGCAAAAAACGCGCAGGCGA 1038  
 258 GTGAAGAGTGGGGCGCTGTGTAACTGCGGGTATTGGCAAAAAAACCGCGCAACGCT 199  
 1039 TCATCCTTGACCTGAAA 1055  
 198 TGATTTGTAATGAAA 182

RESULT 3  
 BH371812  
 LOCUS  
 DEFINITION  
 AG-ND-162110-TR ND-TAM Anopheles gambiae genomic clone  
 AG-ND-162110, genomic survey sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Anopheles gambiae (African malaria mosquito)  
 Anopheles gambiae  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
 Culicidae; Anophelinae; Anopheles.  
 REFERENCE  
 AUTHORS  
 Hong, Y.-S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J.,  
 Ren, C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B.,  
 Gardner, M.J. and Collins, F.H.  
 Construction of a BAC library and generation of BAC end  
 sequence-tagged connectors for genome sequencing of the African  
 malaria mosquito Anopheles gambiae  
 Mol. Genet. Genomics 268 (6), 720-728 (2003)  
 Other GSSEs: AG-ND-162110.TF  
 Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: b.loftus@tigr.org  
 This clone is from an A. gambiae BAC library (ND-TAM) provided by  
 F.H. Collins and sequenced by The Institute for Genomic Research  
 (TIGR). The BAC library was generated from A. gambiae PEST strain  
 DNA. All DNA was extracted from newly hatched first instar larvae  
 to minimize the inclusion of DNA from microorganisms that inhabit  
 the gut. The DNA is derived from mixed sexes of larvae. The BAC  
 library was constructed at Texas A&M University BAC Center  
 University, College Station, Texas 77843-2123, USA using a HindIII  
 partial digest.  
 Seq primer: M13 Rev  
 Class: BAC ends.  
 Location/Qualifiers  
 1..790  
 /organism="Anopheles gambiae"  
 /mol\_type="genomic DNA"  
 /strain="PEST"  
 /db\_xref="taxon:7165"  
 /clone="AG-ND-162110"  
 /clone\_lib="ND-TAM"  
 /note="Vector: pECBAC1; Site\_1: HindIII"

FEATURES  
 source  
 1..790  
 /organism="Anopheles gambiae"  
 /mol\_type="genomic DNA"  
 /strain="PEST"  
 /db\_xref="taxon:7165"  
 /clone="AG-ND-162110"  
 /clone\_lib="ND-TAM"  
 /note="Vector: pECBAC1; Site\_1: HindIII"  
 ORIGIN  
 Query Match 6.1%; Score 64; DB 9; Length 790;



Mon Feb 27 08:33:23 2006

RESULT 6  
CX067473/c  
LOCUS  
DEFINITION 1322269 NCCWA 04RT Oncorhynchus mykiss cDNA, mRNA sequence.  
ACCESSION CX067473  
VERSION CX067473.1 GI:56989039  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Oncorhynchus mykiss (rainbow trout)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
1 (bases 1 to 485)  
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.  
O4RT grool, NCCWA/WVU EST Project, Phase II, in collaboration with  
GRASP  
Unpublished (2004)  
Contact: Rexroad CE  
USDA, ARS, National Center for Cool and Cold Water Aquaculture  
11876 Leetown Road, Kearneysville, WV 25430, USA  
Tel: 304 724 8340 x2129  
Fax: 304 725 0351  
Email: cexroad@nccwa.ars.usda.gov  
Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified with  
cross\_match v0.990329.  
Plate: 106 row: F column: 8  
Seq primer: ATTTAGTGCACATATAG.  
Location/Qualifiers  
1. .485  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/lab\_host="Top10"  
/clone\_lib="NCCWA 04RT"  
/note="Vector: PCR 4-TOPO; This is an early neurogenesis  
SSH library created by Mathew L. Rise constructed by  
subtracting late neurogenesis (mixed stages: hindbrain  
swelling + heart tube with peristalsis) from early  
neurogenesis (mixed stages: neural groove + 1/2 epiboly).  
Fish were from a domesticated strain (Spring Valley Trout  
Farm, Langley, B.C.), courtesy of Bob Devlin, DFO. These  
are mostly internal (coding) sequences."

FEATURES  
source

Query Match 4.5%; Score 47; DB 8; Length 485;  
Best Local Similarity 55.1%; Pred. No. 0.076;  
Matches 92; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 491 AACATGTCATGAAGCGCTAAATCATCCGGAAGCCCAACATCATGTAGAAGATATGAAGA 550  
DB 183 AACATTCATGAAGATGCAATTAATTAAGAATGAGAATGATATGGAATGTATGTG 124  
QY 551 CATGACGCTGTATGCTTTTTCAGCTGAGACAGAGCTTTTTCAGCAACATATGTT 610  
DB 123 CATATGCAAGTTTCTCTTTTATTAACAATGAAGATATTTTGAAGAGGATGTATGTT 64  
QY 611 AACTTTTTCATCTAGCTTTGCTGTTTGTGTACAAATGAAGAGCA 657  
DB 63 AAAGATTGCTTTTAAATGACAGCATCAATTTGATTTATAGAAAGTA 17

ORIGIN

Query Match 4.5%; Score 47; DB 8; Length 485;  
Best Local Similarity 55.1%; Pred. No. 0.076;  
Matches 92; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 491 AACATGTCATGAAGCGCTAAATCATCCGGAAGCCCAACATCATGTAGAAGATATGAAGA 550  
DB 183 AACATTCATGAAGATGCAATTAATTAAGAATGAGAATGATATGGAATGTATGTG 124  
QY 551 CATGACGCTGTATGCTTTTTCAGCTGAGACAGAGCTTTTTCAGCAACATATGTT 610  
DB 123 CATATGCAAGTTTCTCTTTTATTAACAATGAAGATATTTTGAAGAGGATGTATGTT 64  
QY 611 AACTTTTTCATCTAGCTTTGCTGTTTGTGTACAAATGAAGAGCA 657  
DB 63 AAAGATTGCTTTTAAATGACAGCATCAATTTGATTTATAGAAAGTA 17

RESULT 7  
CN384972  
LOCUS  
DEFINITION 745 bp mRNA linear EST 31-AUG-2004  
LE2TR02F22 Tomato CL5915 roots under different developmental stages  
Lycopersicon esculentum cDNA clone LE2TR02F22, mRNA sequence.  
ACCESSION CN384972  
VERSION CN384972.1 GI:51700286  
KEYWORDS EST.  
SOURCE  
ORGANISM  
Lycopersicon esculentum (Solanum lycopersicum)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
AUTHORS  
TITLE

1 (bases 1 to 745)  
Wang, C.K., Chen, P.Y., Wang, H.M., Soong, S.C., Chen, S.C. and To, K.Y.  
DNA microarray profiling of gene expression during tomato root  
development  
Unpublished (2004)

JOURNAL  
COMMENT

Contact: Kin-Ying, To  
Crop Plant Improvement Group  
Institute of Bioagricultural Sciences, Academia Sinica  
128 Academia Rd. Section 2, Taipei, Taiwan 11529  
Tel: 886-2-2653-3161  
Fax: 886-2-2561-5600  
Email: kyto@gate.sinica.edu.tw  
Insert Length: 745 Std Error: 0.00  
Plate: 02 row: F column: 22  
Seq primer: smart2.  
Location/Qualifiers  
1. .745  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="CL5915"  
/db\_xref="taxon:4081"  
/clone="LE2TR02F22"  
/tissue\_type="roots"  
/dev\_stage="1-, 2-, 3-, and 4- month-old"  
/lab\_host="E.coli BM25.8"  
/clone\_lib="Tomato CL5915 roots under different  
developmental stages"  
/note="Vector: pRiplex2; Tomato CL5915 seeds are obtained  
from AVRDC. Roots were harvested from plants grown under  
different developmental stages: 1-, 2-, 3-, 4-month-old.  
Equal aliquots of mRNA of different developmental stages  
were mixed and used for cDNA library construction. (Smart  
PCR cDNA library construction kit, Clontech)"

FEATURES  
source

Query Match 4.3%; Score 45.8; DB 7; Length 745;  
Best Local Similarity 47.6%; Pred. No. 0.19;  
Matches 169; Conservative 0; Mismatches 182; Indels 4; Gaps 1;  
QY 537 AGAAGATATGAAGACATGACGCTGTATGCTTTTTCAGCTGCAGACAGAGCTTTT 596  
DB 358 AGAAGTGGATATATATTGGAAGATAATAGTAGTGGGAAGAAAAAAGAGGATTTT 417  
QY 597 ACAGACATATGTTAACTTTTCTAGCTTTGCTCTTTTGTGTTTCAATGAAGAGC 656  
DB 418 GTCAAAATTTAAATTCCTTAAGAGAAATAGAAGGATGAAGTATTCACACATTCAGNAC 477  
QY 657 AGTCAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGGACGATTTGATGATCG 716  
DB 478 ATGAGAGA- - - - - AAGGGTAATAACAAGATTTAGAAATTAGTCAGCATGATGTTTAC 533  
QY 717 CCCCAATATATTGTCATTGAAACCGCGGATCGGCTATCAGATCTTCACGCCAAATCCG 776  
DB 534 ATTTAATTTTATTTATGCGATCTGTGGATTTTATTTAGTTTAGTTAGCTAGTAG 593  
QY 777 TTTTATTATGAAGAAACAGCAAGAAACAATCTATACATACCATGATTTATGAAGAGAGAC 836  
DB 594 ATTATTAAATTACACAAAATTAAGAAAAGATATATGAAGTATACATATGATGATCTAT 653  
QY 837 ACGAATGCGCTGTACGCGCTTTTCGACAAAGGGAAGAAAAAATGCTGTTTACGAAA 891  
DB 654 TGAATTCATTTGGTGTGCTTGTCAAAAAGAAAAAGCAATTCATTTTACGAAA 708

ORIGIN

Query Match 4.3%; Score 45.8; DB 7; Length 745;  
Best Local Similarity 47.6%; Pred. No. 0.19;  
Matches 169; Conservative 0; Mismatches 182; Indels 4; Gaps 1;  
QY 537 AGAAGATATGAAGACATGACGCTGTATGCTTTTTCAGCTGCAGACAGAGCTTTT 596  
DB 358 AGAAGTGGATATATATTGGAAGATAATAGTAGTGGGAAGAAAAAAGAGGATTTT 417  
QY 597 ACAGACATATGTTAACTTTTCTAGCTTTGCTCTTTTGTGTTTCAATGAAGAGC 656  
DB 418 GTCAAAATTTAAATTCCTTAAGAGAAATAGAAGGATGAAGTATTCACACATTCAGNAC 477  
QY 657 AGTCAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGGACGATTTGATGATCG 716  
DB 478 ATGAGAGA- - - - - AAGGGTAATAACAAGATTTAGAAATTAGTCAGCATGATGTTTAC 533  
QY 717 CCCCAATATATTGTCATTGAAACCGCGGATCGGCTATCAGATCTTCACGCCAAATCCG 776  
DB 534 ATTTAATTTTATTTATGCGATCTGTGGATTTTATTTAGTTTAGTTAGCTAGTAG 593  
QY 777 TTTTATTATGAAGAAACAGCAAGAAACAATCTATACATACCATGATTTATGAAGAGAGAC 836  
DB 594 ATTATTAAATTACACAAAATTAAGAAAAGATATATGAAGTATACATATGATGATCTAT 653  
QY 837 ACGAATGCGCTGTACGCGCTTTTCGACAAAGGGAAGAAAAAATGCTGTTTACGAAA 891  
DB 654 TGAATTCATTTGGTGTGCTTGTCAAAAAGAAAAAGCAATTCATTTTACGAAA 708

RESULT 8  
CNS0039G/c  
LOCUS  
DEFINITION 1101 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR08K10 of RPCL-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
ACCESSION  
AL063921

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.





source

1. 527  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-DUI-aak-p-13-0-UI"  
/tissue\_type="Primary Lung Epithelial Cells"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-DUI"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-DUI is a normalized cDNA library containing the following tissue(s): Primary Lung Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCCTGTAGGC.  
TAG\_SEQ=None found"

# ORIGIN

Query Match 4.1%; Score 42.8; DB 6; Length 527;  
Best Local Similarity 49.1%; Pred. No. 1.2;  
Matches 113; Conservative 0; Mismatches 117; Indels 0; Gaps 0;  
QY 825 GTAAGAGAGACAGAAATGCGCTGACGGCTTTTCGACAAAGGAGAAATGCTGTTT 884  
DB 275 GTGCGGAGGAGCGCCAACTGCTGTTGGCTTCTCAACGGCGCGGAGAAATCCTTC 216  
QY 885 ACGAAATGCTGAATGTTACGGGATCGCCCAAAAGGAGCGCTTGCATCTCGCTTCC 944  
DB 215 CGCGAGCTGATCAAGATCAGCGCGCGTGGCGCGCGCATTCCTCGCTGCTGCTGCG 156  
QY 945 GCGCATCGGAGCGGTGATTGAACGATCGAGAAATGAGCAAGAGCAATTCCTGTCAA 1004  
DB 155 TTGAGTTGAGGAGGATTTGCGGAGCGCTGCGGCGGAGGACACGCGCGCTGCTCAAG 96  
QY 1005 TTTCCCGCGGTAGGCAAAAAACGCAAGGAGATCATCTTGACCTGAA 1054  
DB 95 GTGCGGCGCATCGCAAGAGACGCGCGAGCGCTGCTGCTGCACTCAA 46

RESULT 12  
CX146914/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Rexroad CE  
USDA, ARS, National  
11876 Lestown Road, Kearneysville, WV 25430, USA  
Tel: 304 724 8340 x2129  
Fax: 304 725 0351  
Email: crexroad@cccwa.ars.usda.gov  
Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified with  
cross\_match v0.990329.

TITLE  
Expressed sequence tag analysis of the human pathogen  
Paracoccidioides brasiliensis yeast phase: identification of  
putative homologues of Candida albicans virulence and pathogenicity  
genes  
JOURNAL  
PUBMED  
COMMENT  
Eukaryot. Cell 2 (1), 34-48 (2003)  
12582121  
Contact: Gustavo Henrique Goldman  
Laboratory of Molecular Biology - FCFRP  
Universidade de Sao Paulo - USP - FCFRP  
Av do Cafe S/N, CEP: 14040-903, Ribeirao Preto - SP, Brazil  
Email: ggoldman@usp.br.  
Location/Qualifiers

FEATURES  
source  
1. 519  
/organism="Paracoccidioides brasiliensis"  
/mol\_type="mRNA"  
/db\_xref="taxon:121759"  
/clone\_lib="PB0001"

# ORIGIN

Query Match 4.1%; Score 42.8; DB 5; Length 519;  
Best Local Similarity 55.3%; Pred. No. 1.2;  
Matches 83; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
QY 694 AAGGACGATGATTATGATCGCCCAATATATTGTCATTGAAACGGCGGATCGGCT 753  
DB 145 AAAAGAGGAGGATGAATCAGCCCAACACACACAGCGTCCGCAAGCGGAGAACGGGA 204  
QY 754 ATCAGATCTTCAGCCAAATCGTTTATTAAGAAACACAGCAAGAAACAATCTATA 813  
DB 205 ACCACAATTCATTCAGATCTGATGAAGAAATAGGAAACATGCGAGAGACATCAATT 264  
QY 814 CATACATTTATGTAAGAGAGACACGAATG 843  
DB 265 TATATAGATATGCAAAAATATACGGGGTG 294

RESULT 11  
CB853944/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT  
Contact: McCray, PB  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.reagen.com) or from Open Biosystems  
(www.openbiosystems.com).  
Seq primer: M13 FORWARD  
POLYA=No.  
Location/Qualifiers

```
Plate: 120 row: M column: 2
Seq primer: GTAATACGACTCTACTATAGGG.
Location/Qualifiers
1. .684
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/clone_lib="NCCWA 03RT"
/note="Vector: pBluescript SK+; This is a normalized (Cot
= 5) O. mykiss whole juvenile library created by Matthew
L. Rise from approximately 4 month old O. mykiss
(Tzenaicut Lake strain) obtained from Vancouver Island
Trout Hatchery (Duncan, B.C.)."

ORIGIN
Query Match 4.0%; Score 42.4; DB 8; Length 684;
Best Local Similarity 49.5%; Pred. No. 1.7;
Matches 109; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 177 TGAATTTGAGATAAAGTAGAGCGCTCGCACACTTGAAGACTTTAAAGCTGCTTATAA 236
Db 492 TGGTCTAGATCTAGAGCTTACCCAACTTACCCCAAGGAGGAGGATACATATGCAATYAA 433
QY 237 AGGTGGAGCTCATCGATCAGAAAAAGGGTTATTCTGTTTCGCAACAGGTGACGA 296
Db 432 AAGATGACTGGTTCATTTGGTTCAGAAATATCAGATCAGATTTGTGATGTCAAATTTGGGACCA 373
QY 297 CATTTCTCCCTCAGCAAAACAAACGGTTATATCGGAGTGACTGAAGATGGCGTGATTTC 356
Db 372 AAAACTTATCCATCTGACACAGGCTAAATTTCCAAAGTGTACGCTGAGGCGATTATCAT 313
QY 357 GACTTTTTCACGGTCGCCCGGCGATCTTATCAGAACCCATT 396
Db 312 AACACTTTACCATTTCCCTAGTGTATTGACCTCATATT 273

RESULT 13
CX144590/c 701 bp mRNA linear EST 03-JAN-2005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 701)
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.
03RT rblb, NCCWA/WVU EST Project, Phase II, in collaboration with
GRASP
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: 117 row: G column: 22
Seq primer: GTAATACGACTCTACTATAGGG.
Location/Qualifiers
1. .701
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/lab_host="DH10B"

FEATURES
source
Plate: 120 row: M column: 2
Seq primer: GTAATACGACTCTACTATAGGG.
Location/Qualifiers
1. .684
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/clone_lib="NCCWA 03RT"
/note="Vector: pBluescript SK+; This is a normalized (Cot
= 5) O. mykiss whole juvenile library created by Matthew
L. Rise from approximately 4 month old O. mykiss
(Tzenaicut Lake strain) obtained from Vancouver Island
Trout Hatchery (Duncan, B.C.)."

ORIGIN
Query Match 4.0%; Score 42.4; DB 8; Length 684;
Best Local Similarity 49.5%; Pred. No. 1.7;
Matches 109; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 177 TGAATTTGAGATAAAGTAGAGCGCTCGCACACTTGAAGACTTTAAAGCTGCTTATAA 236
Db 492 TGGTCTAGATCTAGAGCTTACCCAACTTACCCCAAGGAGGAGGATACATATGCAATYAA 433
QY 237 AGGTGGAGCTCATCGATCAGAAAAAGGGTTATTCTGTTTCGCAACAGGTGACGA 296
Db 432 AAGATGACTGGTTCATTTGGTTCAGAAATATCAGATCAGATTTGTGATGTCAAATTTGGGACCA 373
QY 297 CATTTCTCCCTCAGCAAAACAAACGGTTATATCGGAGTGACTGAAGATGGCGTGATTTC 356
Db 372 AAAACTTATCCATCTGACACAGGCTAAATTTCCAAAGTGTACGCTGAGGCGATTATCAT 313
QY 357 GACTTTTTCACGGTCGCCCGGCGATCTTATCAGAACCCATT 396
Db 312 AACACTTTACCATTTCCCTAGTGTATTGACCTCATATT 273

RESULT 14
CX153248/c 738 bp mRNA linear EST 03-JAN-2005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 738)
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.
03RT rblb, NCCWA/WVU EST Project, Phase II, in collaboration with
GRASP
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: 131 row: B column: 19
Seq primer: GTAATACGACTCTACTATAGGG.
Location/Qualifiers
1. .738
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="NCCWA 03RT"
/note="Vector: pBluescript SK+; This is a normalized (Cot
= 5) O. mykiss whole juvenile library created by Matthew
L. Rise from approximately 4 month old O. mykiss
(Tzenaicut Lake strain) obtained from Vancouver Island
Trout Hatchery (Duncan, B.C.)."

ORIGIN
Query Match 4.0%; Score 42.4; DB 8; Length 738;
```

	Best Local Similarity	49.5%;	Pred. No. 1.7;	Mismatches 111;	Indels 0;	Gaps 0;
	Matches 109;	Conservative	0;			
Qy	177	TGGAATTGAGAAATAAAGCTAGAGGCCGTCCGCACACTCGAAGACTTTTAAAGCTGCTTAATAA	236			
Db	492	TGGTGCTAGAATCTAGAGCCTACCCAATTACCCEAGGAGGAGATACATATGCAATAA	433			
Qy	237	AGGDTGSCAGTCATCGCATCAGAAAAAGGGGTTTTATTCTGTTTCGCAAAACAGGTGGACGA	296			
Db	432	AAGNTGACTGGTCANTTGGTCAGATATATCAGATCAGATTGTGATGTCBAATTTGTGGACCA	373			
Qy	297	CATTTCTCCCCTCHGCAAAAACAACGGTTATATCGGAGTAGCTAGAACAGATGGCGTGATTC	356			
Db	372	AAAACTTCATCCCCATCTGAAACAGCTAAAAATCCAAAGTGTACGCTGAAGGGCATTTATCAT	313			
Qy	357	GACTTTTTCAGCGGTGCGCCGGGCATCTTTATCAGAACCCCAATT	396			
Db	312	AACACTTTTACATTTCTCTAGTGTTATTTTGGACCTCATATT	273			

RESULT 15  
AZ674153  
LOCUS  
DEFINITION  
          882 bp DNA linear GSS 14-DEC-2000  
ENTM064TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, genomic survey sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
          GSS.  
          AZ674153  
          AZ674153.1 GI:11811299  
          Entamoeba histolytica  
          Entamoeba histolytica  
          Eukaryota; Entamoebidae; Entamoeba.  
          1 (bases 1 to 882)  
AUTHORS  
          Loftus,B., Van Aken,S. and Fraser,C.  
TITLE  
          Determination of clone end sequences from Entamoeba histolytica  
          HM1:IMSS sheared DNA library  
JOURNAL  
COMMENT  
          Unpublished (2000)  
          Contact: Brendan J Loftus  
          Department of Eukaryotic Genomics  
          The Institute for Genomic Research  
          9712 Medical Center Dr., Rockville, MD 20850, USA  
          Tel: 301 838 0208  
          Fax: 301 838 3543  
          Email: bjloftus@igir.org  
          Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
          DNA library  
          Seq primer: M13-Reverse  
          Class: shotgun  
          High quality sequence start: 17  
          High quality sequence stop: 647.

**FEATURES**  
**SOURCE**

## ORIGIN

Query Match 3.9%; Score 41.2; DB 9; Length 882;  
Best Local Similarity 51.6%; Pred. No. 4;

Search completed: February 27, 2006, 07:55:28  
Job time : 4979 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:31:46 ; Search time 745 Seconds  
(without alignments)  
9437.920 Million cell updates/sec

Title: US-10-510-408-1  
Perfect score: 1055  
Sequence: 1 ccttaagtgtaaggcaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*
- 14: Geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055	100.0	1055	12	ADN60329
2	1055	100.0	1522	12	ADG32153
3	459	43.5	459	6	ABK75799
4	195.4	18.5	546	6	ABK74325
5	137	13.0	595	6	ABK78681
c	124.8	11.8	110000	6	ABAO3041_15
c	122	11.6	110000	6	ABO69245_15
8	122	11.6	110000	6	ABO67195_1
c	108.2	10.3	110000	9	ADB12064_14
10	107.6	10.2	110000	10	ADP77343_05
11	104.8	9.9	654	9	ADB11501
12	104.8	9.9	654	9	ADB11499
13	103.4	9.8	615	6	ABN91117
14	103.4	9.8	615	13	ADN501069
15	89.8	8.5	609	8	ABZ42153
16	89.8	8.5	14273	2	AAV52173
17	89.8	8.5	110000	10	ABS56454_00
18	89	8.4	591	10	ABX05796
19	89	8.4	609	8	ACF72471

20	89	8.4	7159	2	AAV74613
21	88.2	8.4	609	13	ADR92479
22	88.2	8.4	609	14	AEA56349
c	87.8	8.3	110000	6	ABA90521_22
24	87.4	8.3	594	4	AAH90735
25	87.4	8.3	594	13	ADK45627
26	87.4	8.3	603	14	ADW94411
27	87.4	8.3	618	4	AAH90848
28	87.4	8.3	11451	14	ADW94161
29	84.8	8.0	2121	6	ABQ70788
30	74.4	7.1	696	10	ADH85304
31	72.2	6.8	594	6	ABN67714
c	71.8	6.8	110000	6	ABN71527_20
c	70.2	6.7	43539	13	ADV87729
c	70.2	6.7	43539	13	ADV78982
c	70.2	6.7	110000	13	ADV81204_21
36	70	6.6	588	6	ABN67713
37	68.4	6.5	591	13	ADV85324
38	65.6	6.2	633	11	ABD01183
39	65.6	6.2	10996	2	AAK13139
40	65.6	6.2	10996	6	ABS98934
41	61.4	5.8	624	12	ADL02869
42	61.4	5.8	45613	4	AAF28535
43	60.2	5.7	609	9	ADA30941
44	60.2	5.7	657	10	ADF00076
45	56	5.3	618	10	ACF68290

ALIGNMENTS

RESULT 1

ADN60329

ID ADN60329 standard; DNA; 1055 BP.

XX

AC ADN60329;

XX

DT 01-JUL-2004 (first entry)

XX

DE B. licheniformis sporulation related polynucleotide, seq id 1.

XX

Mutant host cell; sporulation; oxidoreductase; transferase; hydrolase;

KW lyase; isomerase; ligase; gene; ds.

XX

OS Bacillus licheniformis.

XX

PN WO2003087148-A2.

XX

PD 23-OCT-2003.

XX

PF 25-MAR-2003; 2003WO-DK000200.

XX

PR 10-APR-2002; 2002DK-00000533.

XX

PA (NOVO ) NOVOZYMES AS.

XX

PI Andersen JT, Jorgensen ST, Rasmussen MD, Olsen PB, Clausen IG;

XX

DR WPI; 2004-122131/12.

XX

PT P-PSDB; ADN60330.

XX

PT A Bacillus licheniformis mutant host cell for producing a product of

XX

PS interest e.g. vitamins, antibiotics and enzymes.

XX

PS Disclosure; SEQ ID NO 1; 319pp; English.

XX

The invention relates to a Bacillus licheniformis mutant host cell derived from a parent B. licheniformis host cell. The mutant host cell is mutated in one or more genes encoding one or more polypeptides involved in sporulation. The host cell comprises one or more heterologous genes present in at least two copies, encoding one or more heterologous polypeptides. The heterologous genes are stably integrated into the genome of the cell without leaving any antibiotic resistance marker genes

at the site of integration. The heterologous genes are transcribed from a heterologous promoter or from an artificial promoter. The heterologous operon, preferably a polyclonal operon. The heterologous polypeptide is an antimicrobial peptide, or a fusion peptide comprising a peptide part which in its native form has antimicrobial activity. The heterologous polypeptide is an enzyme, preferably a secreted enzyme. The enzyme is an enzyme of a class selected from the group of enzyme classes consisting of oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC 3), lyases (EC 4), isomerases (EC 5), and ligases (EC 6). The *Bacillus licheniformis* is useful in a process for producing at least one product of interest, comprising cultivating a *B. licheniformis* mutant host cell in a suitable medium, whereby the said product is produced. The process further comprises isolating or purifying the product of interest. The current sequence represents a *B. licheniformis* sporulation related polynucleotide.

Sequence 1055 BP; 325 A; 208 C; 259 G; 263 T; 0 U; 0 Other;

Query Match 100.0%; Score 1055; DB 12; Length 1055;  
Best Local Similarity 100.0%; Pred. No. 9.8e-314;  
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTTAAGTAAGCGCAAAAAGAGGTGATATGATGACGCGCAAGCAAGTTCAAAATC 60
DB 1 CCTTAAGTAAGCGCAAAAAGAGGTGATATGATGACGCGCAAGCAAGTTCAAAATC 60
QY 61 GGTATTATGCTTATGGAAGTCTGTCGGCGGCTCAGCTTTCACCTGGAGGCTTGCC 120
DB 61 GGTATTATGCTTATGGAAGTCTGTCGGCGGCTCAGCTTTCACCTGGAGGCTTGCC 120
QY 121 GAAAAGCGGCTTAAAGTTCAATCAGCTTGAAGAAGTTTATCTGGACGAGACGTTGGA 180
DB 121 GAAAAGCGGCTTAAAGTTCAATCAGCTTGAAGAAGTTTATCTGGACGAGACGTTGGA 180
QY 181 ATTGAGAATAAAGTAGAGCGGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAAGG 240
DB 181 ATTGAGAATAAAGTAGAGCGGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAAGG 240
QY 241 TGGGAGCTCATCGATCAGAAAAGGGGTTTATCTGTTTCGCAACAGGTGACGACATT 300
DB 241 TGGGAGCTCATCGATCAGAAAAGGGGTTTATCTGTTTCGCAACAGGTGACGACATT 300
QY 301 TCTCCCTCAGCAAAAACAAAGGTTATCGAGTGAAGTGGGCTGATTCGACT 360
DB 301 TCTCCCTCAGCAAAAACAAAGGTTATCGAGTGAAGTGGGCTGATTCGACT 360
QY 361 TTTCAGGTCGCGGGGATCTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 420
DB 361 TTTCAGGTCGCGGGGATCTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 420
QY 421 AAGCGGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACGCGACGAAA 480
DB 421 AAGCGGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACGCGACGAAA 480
QY 481 AAGGAATTGAACATGTCATTGAAGCGGTAAATCATCGGAAGCAACATCATGTAGAA 540
DB 481 AAGGAATTGAACATGTCATTGAAGCGGTAAATCATCGGAAGCAACATCATGTAGAA 540
QY 541 GATATGAAGACATGACGCTGTTATGCTTTTTTTCAGCTGCGACAGAGCTTTTTAGCG 600
DB 541 GATATGAAGACATGACGCTGTTATGCTTTTTTTCAGCTGCGACAGAGCTTTTTAGCG 600
QY 601 AACATATGTTAATCTTTTCATCTAGCTTTGCTGTTTGTGTTACATGAAGACGATC 660
DB 601 AACATATGTTAATCTTTTCATCTAGCTTTGCTGTTTGTGTTACATGAAGACGATC 660
QY 661 AAGAGGTTGAATGAACGTTGATCGAATTCGTAAGGACGATTCATTATGATCGCCCC 720
DB 661 AAGAGGTTGAATGAACGTTGATCGAATTCGTAAGGACGATTCATTATGATCGCCCC 720
QY 721 AATATATGTTCAATGAAAACGCGGATCGGCTATCAGATCTTCACGCCAAATCGGTTTA 780
DB 721 AATATATGTTCAATGAAAACGCGGATCGGCTATCAGATCTTCACGCCAAATCGGTTTA 780
```

```
QY 781 TTTATAAGAAAAACAGCAAAAGAAACAATCTATACATACCATTTATGTAAAGAAACACGA 840
DB 781 TTTATAAGAAAAACAGCAAAAGAAACAATCTATACATACCATTTATGTAAAGAAACACGA 840
QY 841 ATGCGCTGTACGGCTTTTCGACAAAGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATG 900
DB 841 ATGCGCTGTACGGCTTTTCGACAAAGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATG 900
QY 901 TTACGGGATCGGCCCCAAAAGAGGCGCTTGGGATCTCGCTTCGGGCGATCCGGAGCGG 960
DB 901 TTACGGGATCGGCCCCAAAAGAGGCGCTTGGGATCTCGCTTCGGGCGATCCGGAGCGG 960
QY 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGGCGTAGGCA 1020
DB 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGGCGTAGGCA 1020
QY 1021 AAAAAACGGCAAGGACAGATCATCTTCACCTGAAA 1055
DB 1021 AAAAAACGGCAAGGACAGATCATCTTCACCTGAAA 1055

RESULT 2
ADG32153
ID ADG32153 standard; DNA; 1522 BP.
XX
AC ADG32153;
XX
DT 26-FEB-2004 (first entry)
XX
DE DNA encoding a mutant B. licheniformis secreted polypeptide SeqID 123.
XX
KW mutant; host cell; production yield; shelf life; product stability;
XX
OS purity; secreted; gene; ds.
XX
OS Synthetic.
XX
OS Bacillus licheniformis.
XX
PN WO2003093453-A2.
XX
PD 13-NOV-2003.
XX
PF 25-MAR-2003; 2003WO-DK000198.
XX
PR 10-APR-2002; 2002DK-00000534.
XX
PA (NOVO ) NOVOZYMES AS.
XX
PI Andersen JT, Jorgensen ST, Rasmussen MD, Olsen PB, Clausen IG;
XX
WPI; 2004-053045/05.
XX
P-PSDB; ADG32154.
XX
New mutant Bacillus licheniformis host cell secreting 5 % less of one or
more secreted polypeptides than the parent host cell, useful for
producing a product of interest e.g. polypeptides, amino acids or
carbohydrates.
XX
Disclosure; SEQ ID NO 123; 422pp; English.
XX
This invention relates to a novel Bacillus licheniformis (B. licheniformis) mutant host cell derived from a parent B. licheniformis host cell that is mutated in genes encoding secreted polypeptides. Specifically, it refers to the generation of an improved Bacillus host that reduces the need for product purification caused by contaminant secreted native polypeptides in the culture medium. Accordingly, the present invention describes reducing the expression of these native proteins (e.g. proteolytic enzymes, nutrient uptake factors and signal molecules), which in turn makes it easier to purify the heterologous product of interest and therefore improving the production process. Further benefits of a mutated host cell include an increase in total production yield and a longevity of shelf life attributable to improved product stability and purity. This polynucleotide is a DNA sequence
```



CC encoding a mutant B. licheniformis secreted polypeptide of the invention.  
XX Sequence 1522 BP; 470 A; 314 C; 384 G; 354 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1055; DB 12; Length 1522;  
Best Local Similarity 100.0%; Pred. No. 9,8e-314;  
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTAAAGTAAAGCAAAAAGAGGTGATATTGATGTACAGCCGAGCAAGTTCAAAATC 60  
DB 468 CCTTAAAGGTAAAGCAAAAAGAGGTGATATTGATGTACAGCCGAGCAAGTTCAAAATC 527  
QY 61 GGTTTTATTCCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTCACCTGGAGCCCTTGCC 120  
DB 528 GGTTTTATTCCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTCACCTGGAGCCCTTGCC 587  
QY 121 GAAAAGCCGGCTAAAGTTCAAAATCAGACTTGAAGTTCGAGAGGTTTATCTGACGAGACGTTGA 180  
DB 588 GAAAAGCCGGCTAAAGTTCAAAATCAGACTTGAAGTTCGAGAGGTTTATCTGACGAGACGTTGA 647  
QY 181 ATTGAGATTAAGTAGAGCCGCTCGCACACTGGAAGACTTTTAAAGCTGCTTTATAAAGGG 240  
DB 648 ATTGAGATTAAGTAGAGCCGCTCGCACACTGGAAGACTTTTAAAGCTGCTTTATAAAGGG 707  
QY 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAAAACAGTGGACGACATT 300  
DB 708 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAAAACAGTGGACGACATT 767  
QY 301 TCTCCCTCAGCAAAAACCGTTATATCGGAGTGACTGGAAGTGGGTGATTTTCGACT 360  
DB 768 TCTCCCTCAGCAAAAACCGTTATATCGGAGTGACTGGAAGTGGGTGATTTTCGACT 827  
QY 361 TTTACGCTGCGCCGGGCACTTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 420  
DB 828 TTTACGCTGCGCCGGGCACTTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 887  
QY 421 AAGCGCTGGAAGCCGATGCGGATGATCTGCGCAAAAGGATACCATCCGACGAAA 480  
DB 888 AAGCGCTGGAAGCCGATGCGGATGATCTGCGCAAAAGGATACCATCCGACGAAA 947  
QY 481 AAGCAATTTGAACATGTCTAAGCCGTAAATCATCCGAAGCCAAACATCATGTAGAA 540  
DB 948 AAGCAATTTGAACATGTCTAAGCCGTAAATCATCCGAAGCCAAACATCATGTAGAA 1007  
QY 541 GATATGAAGACATGACGCTGTTATGCTTTTTCAGCTGCGACAGAAAGCTTTTTAGCG 600  
DB 1008 GATATGAAGACATGACGCTGTTATGCTTTTTCAGCTGCGACAGAAAGCTTTTTAGCG 1067  
QY 601 AACATATGTTAACTTTTCATTTCTAGCTTTGCTGTTTGTGTTTCAATGAAGACGATC 660  
DB 1068 AACATATGTTAACTTTTCATTTCTAGCTTTGCTGTTTGTGTTTCAATGAAGACGATC 1127  
QY 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATTGATTATGTCGCCCC 720  
DB 1128 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATTGATTATGTCGCCCC 1187  
QY 721 AATATATTGTCTAATGAACGCGGGATCGGCTATCAGATCTTCAAGCCAAATCCGTTTA 780  
DB 1188 AATATATTGTCTAATGAACGCGGGATCGGCTATCAGATCTTCAAGCCAAATCCGTTTA 1247  
QY 781 TTTTAAAGAAAACAGCAAGAAACATCTATACATCATATGTAAGAGAACACGA 840  
DB 1248 TTTTAAAGAAAACAGCAAGAAACATCTATACATCATATGTAAGAGAACACGA 1307  
QY 841 ATGCGCTGTCGGCTTTTCGACAGGGAAGAAAATCTGTTTACGAAAATGCTGAATG 900  
DB 1308 ATGCGCTGTCGGCTTTTCGACAGGGAAGAAAATCTGTTTACGAAAATGCTGAATG 1367  
QY 901 TTACGGGATCGGCCCAAAAGAGCGCTTGGGATCCTCGCTTCCGGCGATCCGGAGCGG 960  
DB 1368 TTACGGGATCGGCCCAAAAGAGCGCTTGGGATCCTCGCTTCCGGCGATCCGGAGCGG 1427  
QY 961 TGATTGAAGCGATCGAGAATGAGGACGAGCATTTCTCGTCAAAATTTCCCGCGGTAGGCA 1020

DB 1428 TGATTGAAGCGATCGAGAATGAGGACGAGCATTTCTCGTCAAAATTTCCCGCGTAGGCA 1487  
QY 1021 AAAAAGCGCAAGGAGGAGATCATCTTGGACCTGAAA 1055  
DB 1488 AAAAAGCGCAAGGAGGAGATCATCTTGGACCTGAAA 1522

RESULT 3  
ABK75799  
ID ABK75799 standard; DNA; 459 BP.  
XX AC ABK75799;  
XX DT 13-AUG-2002 (first entry)  
XX DE Bacillus licheniformis genomic sequence tag (GST) #3090.  
XX KW Differential gene expression; genomic sequenced tag; GST;  
KW altered culture condition; environmental stress;  
KW physiological provocation; ds.  
XX OS Bacillus licheniformis.  
XX PN WO200229113-A2.  
XX PD 11-APR-2002.  
XX PF 05-OCT-2001; 2001WO-US031437.  
XX PR 06-OCT-2000; 2000US-00680598.  
PR 27-MAR-2001; 2001US-0279526P.  
XX PA (NOVO ) NOVOZYMES BIOTECH INC.  
PA (NOVO ) NOVOZYMES AS.  
XX PI Berka R, Clausen IG;  
XX WPI; 2002-416584/44.  
XX PT Monitoring differential expression of several genes in first Bacillus cell relative to expression of same genes in one or more second Bacillus cells, by using substrate containing Bacillus genomic sequenced tag array.  
PS Claim 4; SEQ ID NO 3090; 200pp; English.  
XX The invention describes a method of monitoring differential expression of genes in a first Bacillus cell relative to expression of the genes in other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which Bacillus cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 459 BP; 132 A; 97 C; 118 G; 112 T; 0 U; 0 Other;  
Query Match 43.5%; Score 459; DB 6; Length 459;  
Best Local Similarity 100.0%; Pred. No. 2.2e-130;

Matches	459;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	65	TATTGCTTATTGGAAGTCTGTCGCGCGCTCAGCTTTACCTTGGAGGCTTTGGCCGAAA	124						
Db	1	TATTGCTTATTGGAAGTCTGTCGCGCGCTCAGCTTTACCTTGGAGGCTTTGGCCGAAA	60						
Qy	125	AGCCGGCTAAAGTTCCNAATCAGCTTGGAAAGGTTTATCTGGACGGAGACGTTGGAAATTG	184						
Db	61	AGCCGGCTTAAAGTTCCNAATCAGCTTGGAAAGGTTTATCTGGACGGAGACGTTGGAAATTG	120						
Qy	185	AGAAATAAGTAGAGCGCCCTCGCACACTTGGAAAGACTTTAAAGCTCTTTATAAGGGTGGC	244						
Db	121	AGAAATAAGTAGAGCGCCCTCGCACACTTGGAAAGACTTTAAAGCTCTTTATAAGGGTGGC	180						
Qy	245	AGCTCATCGATCAGAAAAAGGGGTTTATCTGTGTTCCGAAAACAGGTGGACGACATTTCTC	304						
Db	181	AGCTCATCGATCAGAAAAAGGGGTTTATCTGTGTTCCGAAAACAGGTGGACGACATTTCTC	240						
Qy	305	CCCTCAGCAAAACAAACGGTTTATATCGGAGTGACTGGAAGATGGCGTGATTTTCGACTTTTC	364						
Db	241	CCCTCAGCAAAACAAACGGTTTATATCGGAGTGACTGGAAGATGGCGTGATTTTCGACTTTTC	300						
Qy	365	ACGGTCGCGCGGGCATCTTATCAGAACCCATTCATTCGTTTTTCAGATTCGATATAAAGC	424						
Db	301	ACGGTCGCGCGGGCATCTTATCAGAACCCATTCATTCGTTTTTCAGATTCGATATAAAGC	360						
Qy	425	GGCTGGAAAAGCCGATGCGCGGATGATCTGCGCAAGGGATACCATACCGCAGCAAAAAAGG	484						
Db	361	GGCTGGAAAAGCCGATGCGCGGATGATCTGCGCAAGGGATACCATACCGCAGCAAAAAAGG	420						
Qy	485	AATTTGAAACATGTGCTATTGAAGCCGTTAAATCATCCGGAA	523						
Db	421	AATTTGAAACATGTGCTATTGAAGCCGTTAAATCATCCGGAA	459						

RESULT 4	
ABK74325	
ID ABK74325	standard; DNA; 546 BP.
XX	
AC ABK74325;	
XX	
DT 13-AUG-2002	(first entry)
XX	
DE Bacillus licheniformis	genomic sequence tag (GSP) #1616.
XX	
KW Differential gene expression;	genomic sequenced tag; GST;
KW altered culture condition;	environmental stress;
KW physiological provocation;	ds.
XX	
OS Bacillus licheniformis.	
XX	
PN WO200229113-A2.	
XX	
PD 11-APR-2002.	
XX	
PF 05-OCT-2001;	2001WO-US031437.
XX	
PP 06-OCT-2000;	2000US-00680598.
XX	
PR 27-MAR-2001;	2001US-0279526P.
XX	
PA (NOVO ) NOVOZYMES BIOTECH INC.	
PA (NOVO ) NOVOZYMES AS.	
XX	
PI Berka R, Clausen IG;	
XX	
DR WPI; 2002-416684/44.	
XX	
PT Monitoring differential expression	of several genes in first Bacillus
PT cell relative to expression	of same genes in one or more second Bacillus
PT cells, by using substrate	containing Bacillus genomic sequenced tag
XX	
PS Claim 4; SEQ ID NO 1616;	200pp; English.

XX The invention describes a method of monitoring differential expression of  
CC genes in a first Bacillus cell relative to expression of the genes in  
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes  
CC isolated from Bacillus cells to a substrate containing array of Bacillus  
CC genomic sequenced tags (GST), examining the array, and determining  
CC relative gene expression by an observed hybridisation reporter signal of  
CC a spot in the array. The method is useful for measuring the expression of  
CC genes in a first Bacillus cell relative to expression of the same genes  
CC in one or more second Bacillus cells. The method is useful for monitoring  
CC global expression of several genes from a Bacillus cell, discovering new  
CC genes, identifying possible functions of unknown open reading frames and  
CC monitoring gene copy number variation and stability. Monitoring changes  
CC in expression of genes may be used to provide a representation of the way  
CC in which Bacillus cells adapt to changes in culture conditions,  
CC environmental stress or other physiological provocation. Extensive follow  
CC up characterisation is unnecessary, when one spot on an array equals one  
CC gene or one open reading frame, since sequence information is available.  
CC This sequence represents a genomic sequence tag (GST) used in the method  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 546 BP; 165 A; 123 C; 141 G; 116 T; 0 U; 1 Other;

RESULT 5	
ABK78681	
ID	ABK78681 standard; DNA; 595 bp.
XX	
AC	ABK78681;
XX	
DT	13-AUG-2002 (first entry)
XX	
DE	Bacillus clausii genomic sequence tag (GST) #1524.
XX	
KW	Differential gene expression; genomic sequenced tag; GST;
KW	altered culture condition; environmental stress;
KW	physiological provocation; ds.
XX	
OS	Bacillus clausii.
XX	
PN	WO200229113-A2.
XX	
PD	11-APR-2002.
XX	
PF	05-OCT-2001; 2001WO-US031437.
XX	
PR	06-OCT-2000; 2000US-00680598.

PR 27-MAR-2001; 2001US-0279526P.  
XX (NOVO ) NOVOZYMES BIOTECH INC.  
PA (NOVO ) NOVOZYMES AS.

XX Berka R, Clausen IG;  
PI WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus  
XX cell relative to expression of same genes in one or more second Bacillus  
XX cells, by using substrate containing Bacillus genomic sequenced tag  
XX array.

PS Claim 11; SEQ ID NO 5972; 200pp; English.

XX The invention describes a method of monitoring differential expression of  
XX genes in a first Bacillus cell relative to expression of the genes in  
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes  
XX isolated from Bacillus cells to a substrate containing array of Bacillus  
XX genomic sequenced tags (GST), examining the array, and determining  
XX relative gene expression by an observed hybridisation reporter signal of  
XX a spot in the array. The method is useful for measuring the expression of  
XX genes in a first Bacillus cell relative to expression of the same genes  
XX in one or more second Bacillus cells. The method is useful for monitoring  
XX global expression of several genes from a Bacillus cell, discovering new  
XX genes, identifying possible functions of unknown open reading frames and  
XX monitoring gene copy number variation and stability. Monitoring changes  
XX in expression of genes may be used to provide a representation of the way  
XX in which Bacillus cells adapt to changes in culture conditions,  
XX environmental stress or other physiological provocation. Extensive follow  
XX up characterisation is unnecessary, when one spot on an array equals one  
XX gene or one open reading frame, since sequence information is available.  
XX This sequence represents a genomic sequence tag (GST) used in the method  
XX of the invention. Note: The sequence data for this patent did not form  
XX part of the printed specification, but was obtained in electronic format  
XX directly from WIP0 at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 595 BP; 183 A; 115 C; 145 G; 152 T; 0 U; 0 Other;

Query Match 13.0%; Score 137; DB 6; Length 595;  
Best Local Similarity 61.9%; Pred. No. 3.8e-31;  
Matches 237; Conservative 0; Mismatches 140; Indels 6; Gaps 1;

QY 679 TGATCGAATTCGTAAAGGAGCGATTGATGTATGTCGCCCAATATATGTCATTGAAA 738  
DB 4 TGATTGATTATTAAGGGAAATTTGGTTGCAGTTGAACCTGCTTATCTCGTGCAGAGA 63  
QY 739 ACGCGGGATCGGCTATCAGATCTTTCACGCCAAATCC-----GTTTATTTATAGAAAA 792  
DB 64 CCAACGGGATTCGGTTCCAAATTTACTGTGCAACCCCTTACCGTTTTTATAAGCGATTAA 123  
QY 793 ACACCAAGAACATCTATACATACCATTTGTAAGAGAGACACGATCGGCTGTACG 852  
DB 124 ATCGGGAGGTAGTATCCCGACACACCACTATGTTTCGGAGAGACAGCCATTTTG 183  
QY 853 GCTTTTCACAAAGGAAGAAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGGATCG 912  
DB 184 GCTTTTACAAAGAGACAGACGCTTGTGTTTGAAGACCTTTTGAATGCTCTGCGCATG 243  
QY 913 CCCCCAAAAGGAGCGCTTCGATCTTCGTTCCGGCGATCCGGGAGCGGTGATTGAAGCGA 972  
DB 244 GCCCAAAAGGAGCGCTCCCATCTTCGTTCTCGGAGAACCAAGAGATATTATCCATGCA 303  
QY 973 TCGAGATGAGGACGACGACATTTCTGTCAAATTTCCCGGGGTAGGCAAAAAAACGGCAA 1032  
DB 304 TTGAACAAGAGATGAGGACATGCTGTTGTCGTTTCCCTGGAGTCGGCAAAAAACAGCA 363  
QY 1033 GGCAGATCATCTTGCACCTGAAA 1055  
DB 364 GACAATCATTTCTTGACTTAAAA 386

RESULT 6

ABA03041\_15/c  
Continuation (16 of 30) of ABA03041 from base 1500001 (Listeria monocytogenes EGD-e geno  
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041

WP	Fragment Name	Begin	End
WP	ABA03041_00	1	110000
WP	ABA03041_01	100001	210000
WP	ABA03041_02	200001	310000
WP	ABA03041_03	300001	410000
WP	ABA03041_04	400001	510000
WP	ABA03041_05	500001	610000
WP	ABA03041_06	600001	710000
WP	ABA03041_07	700001	810000
WP	ABA03041_08	800001	910000
WP	ABA03041_09	900001	1010000
WP	ABA03041_10	1000001	1110000
WP	ABA03041_11	1100001	1210000
WP	ABA03041_12	1200001	1310000
WP	ABA03041_13	1300001	1410000
WP	ABA03041_14	1400001	1510000
WP	ABA03041_15	1500001	1610000
WP	ABA03041_16	1600001	1710000
WP	ABA03041_17	1700001	1810000
WP	ABA03041_18	1800001	1910000
WP	ABA03041_19	1900001	2010000
WP	ABA03041_20	2000001	2110000
WP	ABA03041_21	2100001	2210000
WP	ABA03041_22	2200001	2310000
WP	ABA03041_23	2300001	2410000
WP	ABA03041_24	2400001	2510000
WP	ABA03041_25	2500001	2610000
WP	ABA03041_26	2600001	2710000
WP	ABA03041_27	2700001	2810000
WP	ABA03041_28	2800001	2910000
WP	ABA03041_29	2900001	2944528

Query Match 11.8%; Score 124.8; DB 6; Length 110000;  
Best Local Similarity 58.2%; Pred. No. 3.7e-26;  
Matches 241; Conservative 0; Mismatches 167; Indels 6; Gaps 1;

QY 648 ATGAAGAGCAGCTCAAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATTGAT 707  
DB 69055 ATGACGTGAAATGAAGAGAGAGCGCAATAAATTTGTACGATTATACATAAAGGACCGTTACG 68996  
QY 708 TATGTATCGCCCCCAATATATTGTCATTCGAAAACGCGGGATCGGCTATCAGATCTTCAGG 767  
DB 68995 ACGATTACACCCGGAATATATTGTTGTTGAGCGAGACAAATCGGCTATCAATAAATATACA 68936  
QY 768 CCAATCCGTTTATTTTATAAGA-----AAAAACAGCAAGAAACAAATCTATACATACCAT 821  
DB 68935 GGGAAACCGTTTTCTCTTCAACGACTAGAGGTACAGAGGTACAGAAAGTCTTTTGTATCAG 68876  
QY 822 TATGTAAGAGAGACACGAATCGCTGTACGCTTTTCGACAGGGAAGAAAAATGCTG 881  
DB 68875 CATGTGAGAGAGGATAATATTCTTTATTTGTTTTCAAACAAACAGAAAGACGTTATTTA 68816  
QY 882 TTTTACGAAATGCTGAATGTTTACGGGATCCGCCCAAGAGCGCTTTCGATCTCCGCT 941  
DB 68815 TTCAAAAATTTTGTAGTGTTCGGGCATTCGACCAAAAGCGCACTAGCCATTATTCG 68756  
QY 942 TCCGCGCATCCGGGAGCGGTGATTGAAGCGATCGAGATGAGGACGAGCAAGCATTTCTCCGTC 1001  
DB 68755 TCAGCGATGTCGTTCCACTAATTTTCAGCGATTGATTCGGAAGACGATGTTTATTTAACC 68696  
QY 1002 AAATTTCCCGCGGTAGGCAAAAAAACGGCAAGGACGATCATCTTTCGACCTGAAA 1055  
DB 68695 AAATTTCCGAGTGTGCGTAAAAAAAACAGCTCGCCAAATTTATTTCTTGATTAAAA 68642

RESULT 7

ABQ69245\_15/c  
Continuation (16 of 31) of ABQ69245 from base 1500001 (Listeria innocua DNA sequence #68  
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession ABq69245



QY 743 CGGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTATATAGAAAAA-----CAG 796  
Db 73338 GGGAGTGGGCTACCACTTTTAAATGGCTAACCCCTTCGCGCTGAACGACCGGCTGGGCA 73279  
QY 797 CAAGAACAACATCTATACATACCATTTATAGAGAACACAGATCGCTGTACGGCTT 856  
Db 73278 AGAAGTGAAGATTTTACCTTACCTGAATGTTAGTCAAGACCAAGCTTAGCGCTCTTTGGTTT 73219  
QY 857 TTCGACAAAGGGAAGAAAAATCTGTTTACGAAAAATGCTGAATGTTACGGGATCGGCC 916  
Db 73218 TCCCGCCAGGATGAAGAAAGCTTGTCTTAAATTAATCATGTTCTCAGGTATTGGGCC 73159  
QY 917 AAAAGAGCGCTTGGATCTCGCTTCGCGCGATCCGGAGCGGTGATGAAGCGCATCGA 976  
Db 73158 CAAAAGTGTCTTGGCTATCTCGCCCAATGAAGACCATGCTGGCTTAATCCAGCGCATTTGA 73099  
QY 977 GAATCAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGGTAGCAAAAAAAGCGCAAGCA 1036  
Db 73098 AAATGAAGACCGACAGTTTTCGACAGCTTTCCCGGGGTAGGAAAAAAGACGCTCTCA 73039  
QY 1037 GATCATCTTGAACCTGAAA 1055  
Db 73038 GATTGTTCTGGACCTCAA 73020

## RESULT 10

ADP77343\_05

Continuation (6 of 20) of ADP77343 from base 500001 (Lactic acid bacteria Lactobacillus  
WP Sequence split into 20 fragments LOCUS ADP77343 Accession Adf77343

WP	Fragment Name	Begin	End
WP	ADP77343_00	1	110000
WP	ADP77343_01	100001	210000
WP	ADP77343_02	200001	310000
WP	ADP77343_03	300001	410000
WP	ADP77343_04	400001	510000
WP	ADP77343_05	500001	610000
WP	ADP77343_06	600001	710000
WP	ADP77343_07	700001	810000
WP	ADP77343_08	800001	910000
WP	ADP77343_09	900001	1010000
WP	ADP77343_10	1000001	1110000
WP	ADP77343_11	1100001	1210000
WP	ADP77343_12	1200001	1310000
WP	ADP77343_13	1300001	1410000
WP	ADP77343_14	1400001	1510000
WP	ADP77343_15	1500001	1610000
WP	ADP77343_16	1600001	1710000
WP	ADP77343_17	1700001	1810000
WP	ADP77343_18	1800001	1910000
WP	ADP77343_19	1900001	1983043

Query Match	10.2%	Score	107.6	DB	10	Length	110000
Best Local Similarity	54.6%	Fred. No.	7.3e-21				
Matches	215	Conservative	0	Mismatches	179	Indels	0
Gaps	0						
QY	661	AAAGAGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATTCATTATGTATCGCCCC	720				
Db	11648	AAAGATGATAGTGAATGTTGATATCTCAAGGATAGTACAAAATTTGATCCAG	11707				
QY	721	AATATATGTCATGAAACGCGGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTA	780				
Db	11708	CTTATATGTTCTAGATGTAACGCGAGTTGGCTATAAGATTTCTTGTCCACTCTTATA	11767				
QY	781	TTTATAAGAAACACGAAGAAACAATCTATACATTCATTTATAGAGAGACAGCA	840				
Db	11768	GTTATGTGAATAATCAACCTGCAACATTTAGTTGACAAAGTTTATAGAGATACCGAA	11827				
QY	841	ATGCGCTGTACGGCTTTTCGACAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATG	900				
Db	11828	TTACTTTGTATGTTTTTGTCTTTAGAGATGAAGATTAATTTTAAATTAATTAAGTG	11887				
QY	901	TTACGGGATCGGCCCAAGAGCGCTTGCATCCTCGCTTCGGCGATCCGGAGCGG	960				

Db 11888 TTTCAGGAATTTGGACCTAAGTCTGCTGGCAATTTATGCTGCTGAAGATATGATTTCTT 11947  
QY 961 TGATTTGAAGCGATCAGAAATGAGGACGAAGCATTTTCGTCAAAATTTCCCGCGTAGGCA 1020  
Db 11948 TGGTAGTGCATTTCAAAATGTAAGTAATACTGACTGTTTCCCGGAGTAGGAA 12007  
QY 1021 AAAAAGCGCAAGGACGATCATCTTTGACCTGAA 1054  
Db 12008 AAAAGACTGCTTCCAGATTTGCTTAGACTTGA 12041

## RESULT 11

ADB11501

ID ADB11501 standard; DNA; 654 BP.

XX ADB11501;

XX 20-NOV-2003 (first entry)

DE Alloiooccus otitis antigenic protein encoding DNA SEQ ID NO:5561.

KW Alloiooccus otitis; antigenic protein; immunogenic; immunisation;  
gene therapy; Gram-positive bacterium; infection; gene; ds.

XX Alloiooccus otitis.

XX WO2003048304-A2.

PD 12-JUN-2003.

XX 25-NOV-2002; 2002WO-US036123.

XX 29-NOV-2001; 2001US-0333777P.

PR 18-NOV-2002; 2002US-0426742P.

XX (AMHP ) WYETH HOLDINGS CORP.

XX Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;  
WPI; 2003-505284/47.

XX P-PSDB; ADB11498.

New Alloiooccus otitis polynucleotides and polypeptides, useful for  
treating and diagnosing diseases, drug screening assays and monitoring of  
effects during drug clinical trials.

Claim 7; SEQ ID NO 5561; 1019pp; English.

The present invention describes an isolated polynucleotide (1) of  
Alloiooccus otitis genomic DNA, which encodes an antigenic protein.  
Alloiooccus otitis is a Gram-positive bacterium. Also described: (1)  
an isolated polypeptide that is encoded by the polynucleotide (1); (2) an  
expression vector comprising the novel isolated polynucleotide (1), its  
complement, degenerate variant or fragment; (3) a genetically engineered  
host cell, transformed, transformed or infected with the vector of (2);  
(4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
composition comprising the polypeptide, its complement, biological  
equivalent or fragment, or the polynucleotide that is comprised in the  
expression vector; (6) a pharmaceutical composition comprising the  
polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
of the polypeptides of (1), their biological equivalent or fragment; (8)  
immunising against Alloiooccus otitis by administering to a host the  
immunogenic composition; (9) detecting and/or identifying Alloiooccus  
otitis in the biological sample; (10) a kit comprising a container  
containing the novel polynucleotide, its degenerate variant or fragment,  
or the antibody of (4); and (11) producing a polypeptide by culturing the  
genetically engineered host cell under conditions suitable to produce the  
polypeptide from the culture. (1) can be used in gene therapy. The  
polynucleotides, polypeptides, antibodies and compositions of the present  
invention can be used for treating and diagnosing diseases, drug  
screening assays and monitoring of effects during drug clinical trials.  
The polynucleotides are useful for expressing and detecting Alloiooccus  
otitis. The present sequence encodes an Alloiooccus otitis antigen

```
CC protein from the present invention.
XX
SQ Sequence 654 BP; 192 A; 148 C; 155 G; 159 T; 0 U; 0 Other;

Query Match          9.9%; Score 104.8; DB 9; Length 654;
Best Local Similarity 56.0%; Pred. No. 3.3e-21;
Matches 243; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 628 TTTGCTGTTTGTGTTACATGAAGACAGCTCAAGAGGTGAATGACGTTGATCGAAT 687
Dd 2 TTTATCTGTGTAATAAATGACTATTGGAAGAGGTGGAAGTA---TGTATGAAT 58

QY 688 TCGTAAAGGACGATTGATTATGATATCGCCCCAATATATTGTCATTGAAACCGCGGGA 747
Dd 59 ACATGGTAGGACAGCTGCTGATGTCCAAACCCAGCTACTTGTGTTTACAAGTGGGGGAG 118

QY 748 TCGGCTATCAGATCTTCAGCCAAATCCGTTTATTTATAAGAAAA-----ACAGCAAAAG 801
Dd 119 TGGGCTACCACTTTTAATGGCTAACCCCTTCCGCTGAACGACCGGCTGGGCCAAGAAG 178

QY 802 AAACAACTATACATACATGATTATGTAAGAGACACGAATCCGCTGTACGGCTTTTCGA 861
Dd 179 TGAAGATTTACCTTTACCTGTAATGTTAGTCAAGACCAAGCTTAGCTCTTTGTTTCCCC 238

QY 862 CAAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAG 921
Dd 239 GCCAGATGAAGAAAGCTTGTTCTAAATTAATCAATGCTCAGGTATTGGGCCCAAAA 298

QY 922 GAGCGTTGCGATCCTCGCTTCCGCGATCCGGGAGCGGTGATTAAGCGCATCGAGAATG 981
Dd 299 GTGCTTTGGCTATCTCGCCCAATGAAGACCATGCTGGCTTAATCCAGCGATTGAAAATG 358

QY 982 AGGACGAGCATTTCTGTCATAATTTCCCGGGTAGGCAAAAAACGCAAGGCAGATCA 1041
Dd 359 AAGACCGCAGCTTTTGGACGCTTTTCCCGGGTAGGCAAAAAACAGCTCTCAGATTG 418

QY 1042 TCCTTGACCTGAAA 1055
Dd 419 TTCTGGACCTCAAA 432

RESULT 12
ADBL1499
ID ADBL1499 standard; DNA; 654 BP.
XX
AC ADBL1499;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloiooccus otitis antigenic protein encoding DNA SEQ ID NO:5563.
XX
KW Alloiooccus otitidis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection; gene; ds.
XX
OS Alloiooccus otitis.
XX
PN WO2003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
XX
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;
XX
WPI; 2003-505284/47.
XX
P-PSDB; ADBL1496.
XX
PT New Alloiooccus otitidis polynucleotides and polypeptides, useful for
```

```
PT treating and diagnosing diseases, drug screening assays and monitoring of
XX effects during drug clinical trials.
XX
PS Claim 7; SEQ ID NO 5563; 1019pp; English.
XX
The present invention describes an isolated polynucleotide (I) of
CC Alloiooccus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiooccus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
CC expression vector comprising the novel isolated polynucleotide (I), its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiooccus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (I) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiooccus
CC otitidis. The present sequence encodes an Alloiooccus otitidis antigen
CC protein from the present invention.
XX
SQ Sequence 654 BP; 192 A; 148 C; 155 G; 159 T; 0 U; 0 Other;

Query Match          9.9%; Score 104.8; DB 9; Length 654;
Best Local Similarity 56.0%; Pred. No. 3.3e-21;
Matches 243; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 628 TTTGCTGTTTGTGTTACATGAAGACAGCTCAAGAGGTGAATGACGTTGATCGAAT 687
Dd 2 TTTATCTGTGTAATAAATGACTATTGGAAGAGGTGGAAGTA---TGTATGAAT 58

QY 688 TCGTAAAGGACGATTGATTATGATATCGCCCCAATATATTGTCATTGAAACCGCGGGA 747
Dd 59 ACATGGTAGGACAGCTGCTGATGTCCAAACCCAGCTACTTGTGTTTACAAGTGGGGGAG 118

QY 748 TCGGCTATCAGATCTTCAGCCAAATCCGTTTATTTATAAGAAAA-----ACAGCAAAAG 801
Dd 119 TGGGCTACCACTTTTAATGGCTAACCCCTTCCGCTGAACGACCGGCTGGGCCAAGAAG 178

QY 802 AAACAACTATACATACATGATTATGTAAGAGACACGAATCCGCTGTACGGCTTTTCGA 861
Dd 179 TGAAGATTTACCTTTACCTGTAATGTTAGTCAAGACCAAGCTTAGCTCTTTGTTTCCCC 238

QY 862 CAAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCCAAAAG 921
Dd 239 GCCAGATGAAGAAAGCTTGTTCTAAATTAATCAATGCTCAGGTATTGGGCCCAAAA 298

QY 922 GAGCGTTGCGATCCTCGCTTCCGCGATCCGGGAGCGGTGATTAAGCGCATCGAGAATG 981
Dd 299 GTGCTTTGGCTATCTCGCCCAATGAAGACCATGCTGGCTTAATCCAGCGATTGAAAATG 358

QY 982 AGGACGAGCATTTCTGTCATAATTTCCCGGGTAGGCAAAAAACGCAAGGCAGATCA 1041
Dd 359 AAGACCGCAGCTTTTGGACGCTTTTCCCGGGTAGGCAAAAAACAGCTCTCAGATTG 418

QY 1042 TCCTTGACCTGAAA 1055
Dd 419 TTCTGGACCTCAAA 432

RESULT 13
ABN91317
```







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:32:27 ; Search time 1034 Seconds  
(without alignments)  
8437.325 Million cell updates/sec

Title: US-10-510-408-1  
Perfect score: 1055  
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacgtgaaa 1055

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA\_Main:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055	100.0	1055	9	US-10-510-408-1
2	459	43.5	459	3	US-09-974-300-3090
3	195.4	18.5	546	3	US-09-974-300-1616
4	137	13.0	595	3	US-09-974-300-5972
5	122	11.6	495269	7	US-10-398-221-8
6	122	11.6	3011208	7	US-10-398-221-2058
7	108.2	10.3	1754382	9	US-10-501-282-6651
8	104.8	9.9	654	9	US-10-501-282-5561
9	104.8	9.9	654	9	US-10-501-282-5563
10	103.4	9.8	615	7	US-10-472-972A-364
11	89.8	8.5	609	7	US-10-474-776-7
12	89.8	8.5	14273	2	US-08-961-527-40
13	89.8	8.5	14273	7	US-10-158-844-40
14	89.8	8.5	2162598	8	US-10-472-928-4979
15	89	8.4	591	8	US-10-472-928-167
16	89	8.4	7159	2	US-08-781-986A-302
17	89	8.4	7159	7	US-10-329-624-302
18	88.2	8.4	609	9	US-10-617-320-1114
19	87.4	8.3	603	8	US-10-857-625-351
20	87.4	8.3	11451	8	US-10-857-625-95
21	84.8	8.0	2121	7	US-10-398-221-3601
22	65.6	6.2	10996	3	US-09-070-927A-202
23	61.4	5.8	45613	7	US-10-672-787-22

ALIGNMENTS

RESULT 1

US-10-510-408-1  
; Sequence 1, Application US/10510408  
; Publication NO. US20050221423A1  
; GENERAL INFORMATION:  
; APPLICANT: Jorgensen, Steen Troels  
; APPLICANT: Olsen, Peter Bjarke  
; APPLICANT: Andersen, Jens Tonne  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Improved Bacillus Host Cell  
; FILE REFERENCE: 10295.204-US  
; CURRENT APPLICATION NUMBER: US/10/510,408  
; CURRENT FILING DATE: 2004-10-05  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 1055  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (34)..(552)  
US-10-510-408-1  
Query Match 100.0%; Score 1055; DB 9; Length 1055;  
Best Local Similarity 100.0%; Pred. No. 3.6e-304; Mismatches 0; Indels 0; Gaps 0;  
Matches 1055; Conservative 0;  
QY 1 CCTTAAGTAAGGCAAAAAAGAGGTGATATTGATGTACAGCGGAAGCAAGTTCAAAATC 60  
DB 1 CCTTAAGTAAGGCAAAAAAGAGGTGATATTGATGTACAGCGGAAGCAAGTTCAAAATC 60  
QY 61 GGTATTGCTTTATTTGGAAGTCTGTGGCGCGCTTCACTGAGGACGCTTGA 180  
DB 61 GGTATTGCTTTATTTGGAAGTCTGTGGCGCGCTTCACTGAGGACGCTTGA 180  
QY 121 GAAAGCGGCTTAAAGTTCAATCCAGCTTGAAGAGTTTATCTGAGCGGACGCTTGA 180  
DB 121 GAAAGCGGCTTAAAGTTCAATCCAGCTTGAAGAGTTTATCTGAGCGGACGCTTGA 180  
QY 181 ATTGAGATAAAGTACAGGCGGCTCGCACACTGGAAGACTTTAAAGCTCTTTATAAGGG 240  
DB 181 ATTGAGATAAAGTACAGGCGGCTCGCACACTGGAAGACTTTAAAGCTCTTTATAAGGG 240  
QY 241 TGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAAAAGCTGGAACGATTT 300  
DB 241 TGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAAAAGCTGGAACGATTT 300

Db 241 TGGAGCTCATCGATCAGAAAAAGGGTTTATTCTGTTTCGCAACAGCTGGACGACATT 300  
Qy 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGAGTGACTGAAGATGGCGTGATTTGCACT 360  
Db 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGAGTGACTGAAGATGGCGTGATTTGCACT 360  
Qy 361 TTTCAGCGTGGCGGGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATA 420  
Db 361 TTTCAGCGTGGCGGGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATA 420  
Qy 421 AAGCGGTGGAAAGCGGATGGCGATGATCTGCGCAAGGATACCATACCGCACGAAA 480  
Db 421 AAGCGGTGGAAAGCGGATGGCGATGATCTGCGCAAGGATACCATACCGCACGAAA 480  
Qy 481 AAGGAATTGAACATGTCTTGAAGCGGTAAATCATCGGNAGCAACATCATGTAGAA 540  
Db 481 AAGGAATTGAACATGTCTTGAAGCGGTAAATCATCGGNAGCAACATCATGTAGAA 540  
Qy 541 GATATGAAGACATGACGCTGTTATGCTCTTTTTCAGCTGCAGACAGAGCTTTTTCAGG 600  
Db 541 GATATGAAGACATGACGCTGTTATGCTCTTTTTCAGCTGCAGACAGAGCTTTTTCAGG 600  
Qy 601 AACATATGTTAACTTTTTCATCTAGCTTTGCTGTTTGTGTTTACAATGAAGACAGTC 660  
Db 601 AACATATGTTAACTTTTTCATCTAGCTTTGCTGTTTGTGTTTACAATGAAGACAGTC 660  
Qy 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGAGACGATTGATGTATCGCCCC 720  
Db 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGAGACGATTGATGTATCGCCCC 720  
Qy 721 AATATATTGTCAATGAAACGCGGGATCGGCTATCAGATCTTCAAGCCAAATCCGTTTA 780  
Db 721 AATATATTGTCAATGAAACGCGGGATCGGCTATCAGATCTTCAAGCCAAATCCGTTTA 780  
Qy 781 TTTATAAGAAAAACAGCAAGAAACAACTATACATACATTATGTAAAGAGAACACGA 840  
Db 781 TTTATAAGAAAAACAGCAAGAAACAACTATACATACATTATGTAAAGAGAACACGA 840  
Qy 841 ATGCCTGTACGGCTTTTCGACAGGGGAAGAAAAATGTTTACGAAAAATGCTGAATG 900  
Db 841 ATGCCTGTACGGCTTTTCGACAGGGGAAGAAAAATGTTTACGAAAAATGCTGAATG 900  
Qy 901 TTACGGGATCGGCGCCCAAGAGCGCTTTCGATCTCGCTTCGGCGATCCGGAGCGG 960  
Db 901 TTACGGGATCGGCGCCCAAGAGCGCTTTCGATCTCGCTTCGGCGATCCGGAGCGG 960  
Qy 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGTAGGCA 1020  
Db 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGTAGGCA 1020  
Qy 1021 AAAAAACGGCAAGGAGATCATCTCTTGACCTGAAA 1055  
Db 1021 AAAAAACGGCAAGGAGATCATCTCTTGACCTGAAA 1055

## RESULT 2

US-09-974-300-3090  
; Sequence 3090, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3090  
; LENGTH: 459  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-3090  
Query Match 43.5%; Score 459; DB 3; Length 459;  
Best Local Similarity 100.0%; Pred. No. 3.7e-126; Indels 0; Gaps 0;  
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 65 TATTGCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTTACCTGAGAGCCCTTTGGCGAAA 124  
Db 1 TATTGCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTTACCTGAGAGCCCTTTGGCGAAA 60  
Qy 125 AGCGGCTTAAAGTTCAAAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 184  
Db 61 AGCGGCTTAAAGTTCAAAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 120  
Qy 185 AGAATAAAGTAGAGCGCGCTCGACACTGGAAGACTTTTAAAGCTCTTTATAAAGGTGGC 244  
Db 121 AGAATAAAGTAGAGCGCGCTCGACACTGGAAGACTTTTAAAGCTCTTTATAAAGGTGGC 180  
Qy 245 AGCTCATCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAACAGCTGACGACATTTCTC 304  
Db 181 AGCTCATCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAACAGCTGACGACATTTCTC 240  
Qy 305 CCCTCAGCAAAACAAACGGTTTATTCGAGTGACTGAAGATGGCGTGATTTTCGACTTTTC 364  
Db 241 CCCTCAGCAAAACAAACGGTTTATTCGAGTGACTGAAGATGGCGTGATTTTCGACTTTTC 300  
Qy 365 ACGTTCGCGCGCGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATAAAGC 424  
Db 301 ACGTTCGCGCGCGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATAAAGC 360  
Qy 425 GCCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAAAAAGG 484  
Db 361 GCCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAAAAAGG 420  
Qy 485 AATTGGAACATGTCATTGAAGCCGTAAATCATCCGGAA 523  
Db 421 AATTGGAACATGTCATTGAAGCCGTAAATCATCCGGAA 459

## RESULT 3

US-09-974-300-1616  
; Sequence 1616, Application US/09974300  
; Patent No. US20020146721A1  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; APPLICANT: Clausen, Ib Groth  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1616  
; LENGTH: 546  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(546)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-974-300-1616  
Query Match 18.5%; Score 195.4; DB 3; Length 546;  
Best Local Similarity 85.5%; Pred. No. 2.7e-47;



```
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
; US-10-398-221-2058
```

```
Query Match 11.6%; Score 122; DB 7; Length 3011208;
Best Local Similarity 59.4%; Pred. No. 3.2e-23;
Matches 228; Conservative 0; Mismatches 150; Indels 6; Gaps 1;

QY 678 TTGATCGAATTGTTAAAGGAGCGATTCGATTTATGATCGCCCCCAATATATTGTCTATTGAA 737
Db 1573493 TTGTACGATTACATAAAAGGAATAGTAAACGATTACACAGAAATATATTGTTGTCGAA 1573434

QY 738 AACCGCGGATCGGTATCAGATCTTCAGCCCAATCCGTTTATTATAAGA-----AA 791
Db 1573433 ACAGACAAATCGGCTATCAATAATTACAGAAACCCATTTTCCTTCAACGATTAGAA 1573374

QY 792 AACGCAAGAAACAATCTATACATACCATATTATGAAGAGAGACACGAATGCGCTGTAC 851
Db 1573373 GGTACTGAAGCCAGTCTTTTGTATCAATGTTTCGGGAGACAAATTCCTTGT 1573314

QY 852 GCGTTTTTCGACAGGGAAGAAAAATGCTTTTACGAAATGCTGAATGTTTACGGGGATC 911
Db 1573313 GGTTCCTCAGTCAACAGAAAGCGTTATCTGTTTCAAAAAATATTAAAGCGTTTCGGGTATT 1573254

QY 912 GCCCCAAAAGGCGCTTCGATCTTCGTTCCGGCGATCCGGGAGCGGTGATTTGAACGG 971
Db 1573253 GGGCCAAAAGGTCGCTAGCTATTATAGCTTCTGGAGATGTAGTACCACTTATTACTGCA 1573194

QY 972 ATCGAGAAATGAGGACGAGCATTTCTCGTCAAAATTTCCGGGTAGGCAAAAAACGGCA 1031
Db 1573193 ATCGAATCGAAGATGATGTTTATTGTACTAAATTTCAAGTGTGCTGTAATAAAAAACAGCA 1573134

QY 1032 AGGCAGATCATCTTGACCTGAAA 1055
Db 1573133 CGTCAAAATTATCTTGATTAAAA 1573110
```

```
RESULT 7
US-10-501-282-6651/c
; Sequence 6651, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-18
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6651
```

```
; LENGTH: 1754382
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; US-10-501-282-6651

Query Match 10.3%; Score 108.2; DB 9; Length 1754382;
Best Local Similarity 56.3%; Pred. No. 3.1e-19;
Matches 247; Conservative 0; Mismatches 183; Indels 9; Gaps 2;

QY 623 CTAGCTTTGCTGTTTGTGTTTACAATGAAGAGCAGTCAAGAGGTGAATGAACCTTGAT 682
Db 1473455 CTAGCTTTATCTGTGTTATATAAATGACTATTGGAAAAAGTGGTGAAGTA---TGTA 1473399

QY 683 CGAATTCGTAAGGAGCAGTGTGATTATCTATCGCCCAATATATTGTTCATTGAAAAACGG 742
Db 1473398 TGAATACATGTTAGGAGCAGTCGTGTGATGTCCAACCCAGCTACTTGGTGTTCACAAAGTGGG 1473339

QY 743 CGGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTATTATTATAAGAAAAA-----CAG 796
Db 1473338 GGGAGTGGCTACCACTTTTAATGGCTAACCCCTTCGCCCTGAACGACCGGCTGGGCA 1473279

QY 797 CAAGAAGAAACAATCTATACATACCAATTTATGAAGAGAGACACGAATGCGCTGTACGGCTT 856
Db 1473278 AGAAGTGAAGATTTACCTTTTACCTGAATGTTAGTCAAGACCAGCTTAGCCTCTTTGGTTT 1473219

QY 857 TTCGACAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCC 916
Db 1473218 TCCCGCCAGGATGAAAAAGCTTGTTCCTAAAATTAATCAATGCTCAGGTAATGGGCC 1473159

QY 917 AAAAGGAGCGCTTCGATCTTCGTTCCGGCGATCCGGGAGCGGTGATTTGAAGCGATCGA 976
Db 1473158 CAAAAGTGTCTTGGCTATCTTGGCCCAATGAAGACCATGCTGGCTTAATCCAGGCGATTGA 1473099

QY 977 GAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGCTAGGCAAAAAACGGCAAGCA 1036
Db 1473098 AAATGAAGACGACAGATTTTTCGCAACGCTTTCCGGGGTAGGAAAAAACAAGCCTCTCA 1473039

QY 1037 GATCATCTTCGACCTGAAA 1055
Db 1473038 GATTGTTCTGGACCTCAAA 1473020
```

```
RESULT 8
US-10-501-282-5561
; Sequence 5561, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5561
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (49) .. (651)
; US-10-501-282-5561
```

```
Query Match          9.9%; Score 104.8; DB 9; Length 654;
Best Local Similarity 56.0%; Pred. No. 3.6e-20;
Matches 243; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 628 TTTGCTGTTTGTGTTACAAAGAGCAGTCAAAAGAGGTGAATCAACGTTGATCGAAT 687
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 TTTATCTGTGTTATATAAATGACTATTGGAAAAAGTGGTGAAGTA---TGTATGAAT 58

QY 688 TCGTAAAGGGAGGATGATTGATATGTCGCCCAATATATTGTCATTGAAAAACGGCGGA 747
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59 ACATGGTAGGGACAGTCGTTGATGTCCAACCCAGCTACTTGGTTTACAAAGTGGGGGAG 118

QY 748 TCGGCTATCAGATCTTTCACGCCCAATCGTATTATTTATAAGAAAA-----ACAGCAAG 801
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 TGGGCTACCACTTTTAAATGGCTAACCCCTTCGGCTGAACCGCTGGGCCCAAGAAG 178

QY 802 AAACAATCTATACATACCAATGATTAAGAGAAGACACGAATCGGCTGTACGGCTTTTCA 861
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 TGAAGATTACCTTTACCTGTAATGTAGTCAAGACCAGCTTAGCCTCTTTGGTTTCCCC 238

QY 862 CAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCAAAAG 921
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 GCCAGGATGAAAAAGCTTTGTTTCTAAATTAATCAATGCTCAGGTAATGGGCCAAAA 298

QY 922 GAGCGCTTGGATCTCGCTTCCGGCGATCCGGGAGCGGTGATTTGAAGCGATCGAATG 981
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 GTGCTTTGGCTATCTCGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATCA 1041

QY 982 AGACGGAAGCAATTTCTGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATCA 1041
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 GTGCTTTGGCTATCTCGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATG 358

QY 982 AGACGGAAGCAATTTCTGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATCA 1041
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1042 TCCTTGACCTGAAA 1055

QY 1042 TCCTTGACCTGAAA 1055
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 TTCTGGACCTCAAA 432

RESULT 9
US-10-501-282-5563
; Sequence 5563, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCWHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5563
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22)..(651)
US-10-501-282-5563

Query Match          9.9%; Score 104.8; DB 9; Length 654;
Best Local Similarity 56.0%; Pred. No. 3.6e-20;
Matches 243; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 628 TTTGCTGTTTGTGTTACAAATGAAGAGCAGTCAAAAGAGGTGAATCAACGTTGATCGAAT 687
```

```
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 TTTATCTGTGTTATATAAATGACTATTGGAAAAAGTGGTGAAGTA---TGTATGAAT 58

QY 688 TCGTAAAGGGAGGATGATTGATATGTCGCCCAATATATTGTCATTGAAAAACGGCGGA 747
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59 ACATGGTAGGGACAGTCGTTGATGTCCAACCCAGCTACTTGGTTTACAAAGTGGGGGAG 118

QY 748 TCGGCTATCAGATCTTTCACGCCCAATCGTATTATTTATAAGAAAA-----ACAGCAAG 801
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 TGGGCTACCACTTTTAAATGGCTAACCCCTTCGGCTGAACCGCTGGGCCCAAGAAG 178

QY 802 AAACAATCTATACATACCAATGATTAAGAGAAGACACGAATCGGCTGTACGGCTTTTCA 861
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 TGAAGATTACCTTTACCTGTAATGTAGTCAAGACCAGCTTAGCCTCTTTGGTTTCCCC 238

QY 862 CAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCAAAAG 921
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 GCCAGGATGAAAAAGCTTTGTTTCTAAATTAATCAATGCTCAGGTAATGGGCCCAAAA 298

QY 922 GAGCGCTTGGATCTCGCTTCCGGCGATCCGGGAGCGGTGATTTGAAGCGATCGAATG 981
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 GTGCTTTGGCTATCTCGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATG 358

QY 982 AGACGGAAGCAATTTCTGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATCA 1041
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 AGACCCGACAGTTTTTTGCAACGCTTTCCGGGGTAGGAAAAAAAACAGCCTCTCAGATTG 418

QY 1042 TCCTTGACCTGAAA 1055
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 TTCTGGACCTCAAA 432

RESULT 10
US-10-724-972A-364
; Sequence 364, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724,972A
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 364
; LENGTH: 615
; TYPE: DNA
; ORGANISM: S. epidermidis
US-10-724-972A-364

Query Match          9.8%; Score 103.4; DB 7; Length 615;
Best Local Similarity 56.4%; Pred. No. 9.2e-20;
Matches 216; Conservative 0; Mismatches 161; Indels 6; Gaps 1;

QY 679 TCATCGAATTCGTAAAGGGAGGATGATTGATATGATCGCCCAATATATTGTCATTGAAA 738
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
14 TGTATGCATATATTAAAGGAACATTATCTCAGTTGTTCCCTACACATGTAGTGGTTGAGA 73

QY 739 ACGCGGGATCGGCTATCAGATCTTTCAGCCAAATTCGTTTATTATTATAAGAAAAACAGCA 798
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 CATGTGGCATAGTTATGAGATACAAAGCCCTAATTCCTATCGTTTCAAATAATCTTGT 133

QY 799 AAGAAACAATC-----TATACATACCATTTATGTAAAGAGAAGACGAATGCGGTGACG 852
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
Db 134 AAAAGAGTCCAAATTTATATCTTCACTAAATTTGTACGAGAGATGCTCAACTACTATATG 193
Qy 853 GCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGGATCG 912
Db 194 GCTTTTAAATGAAGAGGAAAAGAAATGTTTCTTAGCTTAATAAAGTGAAGTGGGATAG 253
Qy 913 GCCCAAGAGGAGCGCTTCGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAAGCGA 972
Db 254 GACCTAAATCAGCTTTAGGGATACCTGCTTCAAGTACACCATGAAGTTAAATGGGCTA 313
Qy 973 TCGAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAACGGCAA 1032
Db 314 TCGAAATGAAATGATGCTTATTTAAACAATTTCTCTGGGATAGAAAGAAACTGCAA 373
Qy 1033 GCAGATCATCTTGAACCTGAAA 1055
Db 374 GACAAATGTTGTAGATTAAAA 396

RESULT 11
US-10-474-776-7
; Sequence 7, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYPEPTIDES
; FILE REFERENCE: AM100649-PT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-7

Query Match 8.5%; Score 89.8; DB 7; Length 609;
Best Local Similarity 54.0%; Pred. No. 1.1e-15;
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

Qy 675 ACCTTGATCGAATTCGTAAGGAGCGATTGATTATGATCGCCCAATATATTCATTT 734
Db 13 ACTATGTACGCATATTTAAAGGAAATCATTAACAAATTTACTGCGCAATATACATTTGTTCTT 72
Qy 735 GAAAACGGCGGATCGGCTAT-----CAGATCTTTCAGCCCAATCCCGTTTATTATAAG 788
Db 73 GAAACCAATGGTATTGGTTATATCTGTCATGTGGCAATCTTATGCGCTATTTCAGGTCAG 132
Qy 789 AAAACAGCAAGAAACAATCTATACATACCAATTTATGTAAGAGAGACACGAATGCGGTG 848
Db 133 GTTAATCAGGAGGCTCAGATTTATGTGTCATCAGGTTGTGCGTGAGGACGCCCATTTGCTT 192
Qy 849 TAGCGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGG 908
Db 193 TATGGATTTTCGTCAGAGGATGAGAAAAGCTCTTTCTTAGTCTGATTTTCGCTCTCTGGG 252
Qy 909 ATCGGCGCAAGAGGAGCGCTTCGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAA 968
Db 253 ATTGGTCTGTATCAGCTCTTGTCTATATCTGCTGCTGCTGATGACATGCTGGCTTGGTTCAA 312
Qy 969 GCGATCGAAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAACG 1028
Db 313 GCAATTTGAAACCAAGAACATCACCTACTTGTACCAAGTTCCCTAAAATTTGCGCAAGAAACA 372
Qy 1029 GCAAGCGATCATCTTTCACCTGAAA 1055
Db 373 GCCACGACATGGTGTGACTTGGAA 399

RESULT 12
US-08-961-527-40
```

```
; Sequence 40, Application US/08961527
; Publication No. US20020032323A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-40

Query Match 8.5%; Score 89.8; DB 2; Length 14273;
Best Local Similarity 54.0%; Pred. No. 6.4e-15;
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

Qy 675 ACCTTGATCGAATTCGTAAGGAGCGATTGATTATGATCGCCCAATATATTCATTT 734
Db 11348 ACTATGTACGCATATTTAAAGGAAATCATTAACCAAAATTTACTGCCAAATACATTTGTTCTT 11407
Qy 735 GAAAACGGCGGATCGGCTAT-----CAGATCTTTCAGCCCAATCCCGTTTATTATAAG 788
Db 11408 GAAACCAATGGTATTGGTTATATCTGTCATGTGGCAATCTTATGCGCTATTTCAGGTCAG 11467
Qy 789 AAAACAGCAAGAAACAATCTATACATACCAATTTATGTAAGAGAGACACGAATGCGGTG 848
Db 11468 GTTAATCAGGAGGCTCAGATTTATGTGTCATCAGGTTGTGCGTGAGGACGCCCATTTGCTT 11527
Qy 849 TAGCGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGG 908
Db 11528 TATGGATTTTCGTCAGAGGATGAGAAAAGCTCTTTCTTAGTCTGATTTTCGCTCTCTGGG 11587
Qy 909 ATCGGCGCAAGAGGAGCGCTTCGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAA 968
Db 11588 ATTGGTCTGTATCAGCTCTTGTCTATATCTGCTGCTGCTGATGACATGCTGGCTTGGTTCAA 11647
Qy 969 GCGATCGAAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAACG 1028
Db 11648 GCCATTGAAACCAAGAACATCACCTACTTGTACCAAGTTCCCTAAAATTTGCGCAAGAAACA 11707
Qy 1029 GCAAGCGATCATCTTTCACCTGAAA 1055
Db 11708 GCCACGACATGGTGTGACTTGGAA 11734
```

## RESULT 13

US-10-158-844-40  
; Sequence 40, Application US/10158844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kueh et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB340P1D1  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14273 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:  
US-10-158-844-40

Query Match 8.5%; Score 89.8; DB 7; Length 14273;

Best Local Similarity 54.0%; Pred. No. 6.4e-15;

Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY	675	ACGTTGATCGAATTCGTAAGGAGCGATTGATTGATCGCCCAATATATTGTCATT	734
Db	11348	ACTATGTACGCATATTTAAAGGAATCAATCCAAATTTACTGCCAAATACATTGTCCT	11407
QY	735	GAAGAGCGGGATCGGCTAT-----CAGATCTTCACGCCAAATCCGTTTATTATAAG	788
Db	11408	GAACCAATGGTATGGTTATATCTCGATGTGGCAATCTTATCGCTATTCAGGTCAG	11467
QY	789	AAAAACAGCAAGAAACAATCTATACATCCATTATGTAAGAGAACACGAAATCGGCTG	848
Db	11468	GTTAATCAGAGGCTCAGATTATATGTCATCAGTTGTGCGTGAGGACGCCATTTGCTT	11527
QY	849	TACGGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAATGCTGAATGTTACGGGG	908
Db	11528	TATGGATTTTCGCTCAGAGGATGAGAAAAGCTCTTTCTAGTCTGATTTCCGGTCTGGG	11587
QY	909	ATCGGCCAAAAGAGCGCTTCGATCCTCGCTTCGGGCGATCCGGGAGCGGTGATTGAA	968
Db	11588	ATTGGTCTGTATCAGCTCTTGCTATTTATCGCTGTGATGACAACTGCTGGTGGTTCAA	11647
QY	969	GCATCGAATGAGGACGAGCATTTCTCGTCAAAATTTCCCGGCGTAGGCAAAAAACG	1028
Db	11648	GCCATTGAAACCAAGAACATCACTACTTGACCAAGTTCCCTAAATTTGCAAGAAAACA	11707
QY	1029	GCAAGCAGATCATCTTGCACCTGAAA	1055
Db	11708	GCCAGCAGATGCTGCTGGACTTGGAA	11734

## RESULT 14

US-10-472-928-4979  
; Sequence 4979, Application US/10472928  
; Publication No. US20050020813A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH  
; FILE REFERENCE: P026926WO  
; CURRENT APPLICATION NUMBER: US/10/472,928  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: GB-0107658.7  
; NUMBER OF SEQ ID NOS: 4979  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 4979  
; LENGTH: 2162598  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-472-928-4979

Query Match 8.5%; Score 89.8; DB 8; Length 2162598;

Best Local Similarity 54.0%; Pred. No. 1.1e-13;

Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY	675	ACGTTGATCGAATTCGTAAGGAGCGATTGATTGATCGCCCAATATATTGTCATT	734
Db	65536	ACTATGTACGCATATTTAAAGGAATCAATCCAAATTTACTGCCAAATACATTGTCCT	65595
QY	735	GAAGAGCGGGATCGGCTAT-----CAGATCTTCACGCCAAATCCGTTTATTATAAG	788
Db	65596	GAACCAATGGTATGGTTATATCTCGATGTGGCAATCTTATCGCTATTCAGGTCAG	65655
QY	789	AAAAACAGCAAGAAACAATCTATACATCCATTATGTAAGAGAACACGAAATCGGCTG	848
Db	65656	GTTAATCAGAGGCTCAGATTATATGTCATCAGTTGTGCGTGAGGACGCCATTTGCTT	65715
QY	849	TACGGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAATGCTGAATGTTACGGGG	908
Db	65716	TATGGATTTTCGCTCAGAGGATGAGAAAAGCTCTTTCTAGTCTGATTTCCGGTCTGGG	65775
QY	909	ATCGGCCAAAAGAGCGCTTCGATCCTCGCTTCGGGCGATCCGGGAGCGGTGATTGAA	968
Db	65776	ATTGGTCTGTATCAGCTCTTGCTATTTATCGCTGTGATGACAACTGCTGGTGGTTCAA	65835
QY	969	GCATCGAATGAGGACGAGCATTTCTCGTCAAAATTTCCCGGCGTAGGCAAAAAACG	1028
Db	65836	GCCATTGAAACCAAGAACATCACTACTTGACCAAGTTCCCTAAATTTGCAAGAAAACA	65895
QY	1029	GCAAGCAGATCATCTTGCACCTGAAA	1055
Db	65896	GCCAGCAGATGCTGCTGGACTTGGAA	65922

## RESULT 15

US-10-472-928-167  
; Sequence 167, Application US/10472928  
; Publication No. US20050020813A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH  
; FILE REFERENCE: P026926WO  
; CURRENT APPLICATION NUMBER: US/10/472,928  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: GB-0107658.7  
; NUMBER OF SEQ ID NOS: 4979  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 167  
; LENGTH: 591

```

; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-472-928-167

Query Match      8.4%; Score 89; DB 8; Length 591;
Best Local Similarity 54.0%; Pred. No. 1.8e-15;
Matches 207; Conservative 0; Mismatches 170; Indels 6; Gaps 1;

Qy 679 TGAATCGAATTCGTAAAGGAGGAGATTGATATGTCGCCCAATATATATGTCATTGAAA 738
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 TGTACGCATATTAAAGGAATCATPACCAAAATTACTGCCAAATACATTGTTCTTGAAA 61
Qy 739 ACGGGGGATCGGCTAT-----CAGATCTTCACGCCCAATCCGTTTATTATTAAGAAA 792
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 CCAATGGTATGTGTTATATCCTGCAATGCGCAATCCTTATGCCTATTCAAGTCAAGTTA 121
Qy 793 ACAGCAAGAAACAATCTATACATACATTATGTAGAGAGACACAGAAATGCGGTACG 852
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 ATCAGGAGGCTCAGATTATGTGCATCAGGTTGTGCGTGAGGACGCCCATTTGCTTTATG 181
Qy 853 GCTTTTCGACAAGGGAGAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGATCG 912
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 GATTCGCTCAGAGGATGAGAAAAGCTCTTCTTAGTCTGATTTGGTCTCTGGGATTG 241
Qy 913 GCCCAAAAGGAGGCTTGGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAAGCGA 972
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 GTCCGTGATCAGCTCTTGCTATATATCGCTGCTGATGACAATGCTGGCTTGGTTCAAGCCA 301
Qy 973 TCGAGAAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGCTAGGCAAAAAAAGCGCAA 1032
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 TTGAACCAAGAACATCACCCTACTTGACCAAGTTCCCTAAATTGGCAAGAAAACAGCCC 361
Qy 1033 GGCAGATCATCCTTGACCTGAAA 1055
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 AGCAGATGGTGTGGACTTGGAA 384
```

Search completed: February 27, 2006, 08:12:58  
Job time : 1047 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:31:38 ; Search time 939 Seconds  
(without alignments)  
2395.790 Million cell updates/sec

Title: US-10-510-408-1  
Perfect score: 1055  
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacgtgaaa 1055

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7209121 seqs, 1066183437 residues

Total number of hits satisfying chosen parameters: 14418242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_New.\*  
1: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
11: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
12: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
13: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1055	100.0	1522	8	US-10-510-386-123 Sequence 123, App
2	48.8	4.6	612	8	US-10-467-657-2695 Sequence 2695, App
3	48.8	4.6	612	8	US-10-467-657-2695 Sequence 6709, App
4	39	3.7	600	12	US-11-098-686-10086 Sequence 10086, A
5	39	3.7	1457619	12	US-11-098-686-8739 Sequence 8739, App
6	38.8	3.7	636	6	US-09-925-065A-102700 Sequence 102700, App
7	38.8	3.7	636	6	US-09-925-065A-102701 Sequence 102701, App
8	37.6	3.6	2436	8	US-10-793-626-4398 Sequence 4398, App
9	36.8	3.5	644	6	US-09-925-065A-295479 Sequence 295479, App
10	36.8	3.5	611587	12	US-11-117-187-209 Sequence 209, App
11	36.6	3.5	364	6	US-09-925-065A-174583 Sequence 174583, App
12	36.6	3.5	575	6	US-09-925-065A-147413 Sequence 147413, App
13	36.4	3.5	580	6	US-09-925-065A-720480 Sequence 720480, App
14	36.4	3.5	581	6	US-09-925-065A-477253 Sequence 477253, App
15	36.4	3.5	2279	12	US-11-091-883-216 Sequence 216, App
16	36.4	3.5	212805	12	US-11-112-908-19 Sequence 19, App
17	36.2	3.4	1095	6	US-09-925-065A-42644 Sequence 42644, A
18	36.2	3.4	2092	8	US-10-750-185-64850 Sequence 64850, A
19	36.2	3.4	2092	8	US-10-750-623-64850 Sequence 64850, A
20	35.8	3.4	1911	8	US-10-750-185-32606 Sequence 32606, A

ALIGNMENTS

RESULT 1

US-10-510-386-123  
; Sequence 123, Application US/10510386  
; Publication No. US2005024922A1  
; GENERAL INFORMATION:  
; APPLICANT: Clausen, Jens Tonne  
; APPLICANT: Andersen, Ib Groth  
; APPLICANT: Jorgensen, Steen Troels  
; APPLICANT: Olsen, Peter Bjarke  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Improved Bacillus Host Cell  
; FILE REFERENCE: 10294.204-US  
; CURRENT APPLICATION NUMBER: US/10/510,386  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 248  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 123  
; LENGTH: 1522  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (501)..(1019)  
; US-10-510-386-123

Query Match 100.0%; Score 1055; DB 8; Length 1522;  
Best Local Similarity 100.0%; Pred. No. 1.3e-276;  
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAAGGTAAAGCAAAAAAGGATGATTTGATGTACAGCGCAAGTTCAAAATC 60  
Db 468 CTTAAGGTAAAGCAAAAAAGGATGATTTGATGTACAGCGCAAGTTCAAAATC 527  
QY 61 GCTTTATGCTTATTGGAAGTCTGTCGGCGGCTCAGCTTCACTGGAGGCTTGGCC 120  
Db 528 GCTTTATGCTTATTGGAAGTCTGTCGGCGGCTCAGCTTCACTGGAGGCTTGGCC 587  
QY 121 GAAAGCCGGCTTAAAGTTCAAAATCAGACTTGAAGAGTTTATCTGCACGACGTTGGA 180  
Db 588 GAAAGCCGGCTTAAAGTTCAAAATCAGACTTGAAGAGTTTATCTGCACGACGTTGGA 647  
QY 181 ATTGAGATAAAGTAGAGCCGCTCGCACACTGGAAGACTTTAAAGCTCTTATAAGGG 240

Db 648 ATTGAGATAAAGTAGAGCCGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAAGG 707  
Qy 241 TGGCAGCTCATCGATCAGAAAAGGGGTTTATCTGTTTCGCAACACAGGTGACGACAT 300  
Db 708 TGGCAGCTCATCGATCAGAAAAGGGGTTTATCTGTTTCGCAACACAGGTGACGACAT 767  
Qy 301 TCTCCCTCAGCAAAACAAAGGTTTATCTGAGTGACTGAGATGGCGTGAATTCGACT 360  
Db 768 TCTCCCTCAGCAAAACAAAGGTTTATCTGAGTGACTGAGATGGCGTGAATTCGACT 827  
Qy 361 TTTACGGTTCGCGCGGATCTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 420  
Db 828 TTTACGGTTCGCGCGGATCTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 887  
Qy 421 AAGCGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGGACGAAA 480  
Db 888 AAGCGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGGACGAAA 947  
Qy 481 AAGGAATTTGAACATGTCAATTCGAGCGTAAATCATCCGGAAGCAACATCATGTAGAA 540  
Db 948 AAGGAATTTGAACATGTCAATTCGAGCGTAAATCATCCGGAAGCAACATCATGTAGAA 1007  
Qy 541 GATATGAAGACATGACGCTGTTATGTCTTTTTTCAGCTGCAGACAGAGCTTTTTAGCG 600  
Db 1008 GATATGAAGACATGACGCTGTTATGTCTTTTTTCAGCTGCAGACAGAGCTTTTTAGCG 1067  
Qy 601 AACATATGTTAACTTTTCAATCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGTTC 660  
Db 1068 AACATATGTTAACTTTTCAATCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGTTC 1127  
Qy 661 AAGAGGTGAATGACGTTGATCGAATTCGTAAGGAGGAGTATGATGATCGCCC 720  
Db 1128 AAGAGGTGAATGACGTTGATCGAATTCGTAAGGAGGAGTATGATGATGATCGCCC 1187  
Qy 721 AATATATTGCTATTGAAAACGGGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTA 780  
Db 1188 AATATATTGCTATTGAAAACGGGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTA 1247  
Qy 781 TTTATAAGAAAACAGCAAGAAACAACTATATACATATACATTTATGTAAGAGAAGACAGA 840  
Db 1248 TTTATAAGAAAACAGCAAGAAACAACTATATACATATACATTTATGTAAGAGAAGACAGA 1307  
Qy 841 ATGCGCTGTACGGCTTTTCGACAGGAGGAGAAAATGCTGTTTACGAAAATGCTGAATG 900  
Db 1308 ATGCGCTGTACGGCTTTTCGACAGGAGGAGAAAATGCTGTTTACGAAAATGCTGAATG 1367  
Qy 901 TTACGGGGATCGGCCCAAGAGCGCTTCGATCTCGCTTCGCGCATCCGGGAGCGG 960  
Db 1368 TTACGGGGATCGGCCCAAGAGCGCTTCGATCTCGCTTCGCGCATCCGGGAGCGG 1427  
Qy 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGTAGGCA 1020  
Db 1428 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGTAGGCA 1487  
Qy 1021 AAAAAACGCAAGCGAGATCATCTTGACCTGAAA 1055  
Db 1488 AAAAAACGCAAGCGAGATCATCTTGACCTGAAA 1522

RESULT 2  
US-10-467-657-2695  
; Sequence 2695, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11

; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 2695  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-2695  
Query Match 4.6%; Score 48.8; DB 8; Length 612;  
Best Local Similarity 49.6%; Pred. No. 0.0032;  
Matches 125; Conservative 0; Mismatches 127; Indels 0; Gaps 0;  
Qy 801 GAAACAATCTATACATACCATTTGTAAGAGAAGACAGATGCGCTGTACGGCTTTTCG 860  
Db 163 GTACAACATGTTTATCCAGCTTATCATTCGGGAAGACGACATCTTTTATTGTTTGGC 222  
Qy 861 ACAAGGGAAGAAAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGATCCGCCCAAAA 920  
Db 223 ACTCGGAAGAACCGAAGACCTTCGCCCAACTGATTAAGTTCGGCGGATCGCGCGAAA 282  
Qy 921 GGAGCGCTTGGGATCTCTCGCTTCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAGAAT 980  
Db 283 ACGGCTTTGGGCAATTTTGTGCGCGATGACGCGAGAGCTGGCGGCGGTTCAGAA 342  
Qy 981 GAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGGTAGGCAAAAAACCGCAAGGCAGATC 1040  
Db 343 GAAGATGTCAAAACGCTCTCTCGCTCCGGGAATCGGCAAAAAACCGCGAAGCTATG 402  
Qy 1041 ATCCTTGACCTG 1052  
Db 403 GTCCTTGGAACTG 414  
RESULT 3  
US-10-467-657-6709  
; Sequence 6709, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO 6709  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-6709  
Query Match 4.6%; Score 48.8; DB 8; Length 612;  
Best Local Similarity 49.6%; Pred. No. 0.0032;  
Matches 125; Conservative 0; Mismatches 127; Indels 0; Gaps 0;  
Qy 801 GAAACAATCTATACATACCATTTGTAAGAGAAGACAGATGCGCTGTACGGCTTTTCG 860  
Db 163 GTACAACATGTTTATCCAGCTTATCATTCGGGAAGACGACATCTTTTATTGTTTGGC 222  
Qy 861 ACAAGGGAAGAAAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGATCCGCCCAAAA 920  
Db 223 ACTCGGAAGAACCGAAGACCTTCGCCCAACTGATTAAGTTCGGCGGATCGCGCGAAA 282  
Qy 921 GGAGCGCTTGGGATCTCTCGCTTCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAGAAT 980

Db 283 ACGGCTTGGCATTGTCGGCGATGACGGCAGACGAGCTGGCGGGGGTTCAGAA 342  
QY 981 GAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAAACGGCAAGCGAGATC 1040  
Db 343 GAAGATGTCAAACGCTCTCTCGCTCCGGGAATCGGCAAAAAACCGCGGAACGTATG 402  
QY 1041 ATCCTTGACCTG 1052  
Db 403 GTCTTGAACCTG 414

RESULT 4  
US-11-098-686-10086  
; Sequence 10086, Application US/11098686  
; Publication No. US20060024696A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: 09531-128001  
; CURRENT APPLICATION NUMBER: US/11/098,686  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: PCT/US03/31318  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: US 60/416,395  
; PRIOR FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 11433  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10086  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Lawsonia intracellularis  
US-11-098-686-10086

Query Match 3.7%; Score 39; DB 12; Length 600;  
Best Local Similarity 46.5%; Pred. No. 1.5; Mismatches 0; Gaps 0;  
Matches 126; Conservative 0; Indels 145; Indels 0; Gaps 0;  
QY 785 TAAGAAAAACAGCAAGAACCAATCTATACATACCACTATGTAAGAGAAAGACACGAATGC 844  
Db 120 TAATAAAGGAGAACAAAGTACGTTTTATATTTGTCTATATGTCAGGAGATATACAAGA 179  
QY 845 GCTGTACGCTTTTCGACAAAGGAGAAAATGCTGTTTACGAAAATGCTGAATGTTAC 904  
Db 180 ACTGTTGGCTTTTGAACATGGATGAGCGACAAACGTTTATTTTAACTATCTATATC 239  
QY 905 GGGGATCGGCCCAAAAGGAGCGCTTCGATCTCTCGTCCGGCGATCCGGGAGCGGTAT 964  
Db 240 TAAAGTCGTCGACGTACGACACTTAACTGCTTTCACTTTTCGACCAAAATGATCTAAG 299  
QY 965 TGAAGCGATCGAAGATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAA 1024  
Db 300 ACAATGTGTTATAGAAGAGATATTTTCTCTTACTCTGCTGATCTGGAATTGGAAGAA 359  
QY 1025 AACGGCAGGCGAGATCATCTTGACCTGAAA 1055  
Db 360 GACAGCAACAATATTTTCTTGAACCTAAA 390

RESULT 5  
US-11-098-686-8739  
; Sequence 8739, Application US/11098686  
; Publication No. US20060024696A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.  
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: 09531-128001  
; CURRENT APPLICATION NUMBER: US/11/098,686  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: PCT/US03/31318  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: US 60/416,395

; PRIOR FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 11433  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8739  
; LENGTH: 1457619  
; TYPE: DNA  
; ORGANISM: Lawsonia intracellularis  
US-11-098-686-8739

Query Match 3.7%; Score 39; DB 12; Length 1457619;  
Best Local Similarity 46.5%; Pred. No. 35; Mismatches 145; Indels 0; Gaps 0;  
Matches 126; Conservative 0; Indels 145; Indels 0; Gaps 0;  
QY 785 TAAGAAAAACAGCAAGAACCAATCTATACATACCACTATGTAAGAGAAAGACACGAATGC 844  
Db 1455272 TAATAAAGGAGAACAAAGTACGTTTTATATTTGTCTATATGTCAGGAGATATACAAGA 1455331  
QY 845 GCTGTACGCTTTTCGACAAAGGAGAAAATGCTGTTTACGAAAATGCTGAATGTTAC 904  
Db 1455332 ACTGTTGGCTTTTGAACATGGATGAGCGACAAACGTTTATTTTAACTATCTATATC 1455391  
QY 905 GGGGATCGGCCCAAAAGGAGCGCTTCGATCTCTCGTCCGGCGATCCGGGAGCGGTAT 964  
Db 1455392 TAAAGTCGTCGACGTACGACACTTAACTGCTTTCACTTTTCGACCAAAATGATCTAAG 1455451  
QY 965 TGAAGCGATCGAAGATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAA 1024  
Db 1455452 ACAATGTGTTATAGAAGAGATATTTTCTCTTACTCTGCTGATCTGGAATTGGAAGAA 1455511  
QY 1025 AACGGCAGGCGAGATCATCTTGACCTGAAA 1055  
Db 1455512 GACAGCAACAATATTTTCTTGAACCTAAA 1455542

RESULT 6  
US-09-925-065A-102700/c  
; Sequence 102700, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 102700  
; LENGTH: 636  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-102700

Query Match 3.7%; Score 38.8; DB 6; Length 636;  
Best Local Similarity 50.5%; Pred. No. 1.7; Mismatches 92; Indels 0; Gaps 0;  
Matches 94; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
QY 489 TGAACATGCTATTGAAGCGCGTAAATCATCCGGAAGCCCAACATCATGTAGAGATATGAA 548  
Db 208 TGCCCATGACTTTTAGAAGTTAAGTCAITCTGCAATATCAAGGTACTCAGCTTCTGCTGAA 149  
QY 549 GACATGACGCTGTTATGCTTTTTCAGCTGCAGACAGAAGCTTTTTTAGCGAACATATG 608

Db	148	TGCTTATCTCTTCTAAGTCTAATTGAGCAATATGAATAATTTTGGATGGAGATCTCATC	89
Qy	609	TTAACTTTTTCATCTAGCTTTGCTGTTTGTGTACAAATGAAGAGCAGTCAAGAGGT	668
Db	88	TTAAATGTTTCCCTCTACCTACTTCTTAAATGTCCTATCAATGCAATGAACAATTT	29
Qy	669	GAATGA	674
Db	28	CCTTGA	23
RESULT 7			
US-09-925-065A-102701/c			
; Sequence 102701, Application US/09925065A			
; Publication No. US20040181048A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, David G.			
; TITLE OF INVENTION: Identification and Mapping of Single			
; FILE REFERENCE: 10827.135			
; CURRENT APPLICATION NUMBER: US/09/925,065A			
; PRIOR FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/243,096			
; PRIOR FILING DATE: 2000-10-24			
; PRIOR APPLICATION NUMBER: US 60/252,147			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: US 60/250,092			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: US 60/261,766			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/289,846			
; PRIOR FILING DATE: 2001-05-09			
; NUMBER OF SEQ ID NOS: 957086			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 102701			
; LENGTH: 636			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-065A-102701			
Query Match			
Best Local Similarity 3.7%; Score 38.8; DB 6; Length 636;			
Matches 94; Conservative 0; Mismatches 92; Indels 0; Gaps 0;			
Qy	489	TGAACATGTCATTGAAGCGGTAAATCAATCCGGAACCAACATCATGTAGAGATGAA	548
Db	208	TGCCCATGACTTTTGAAGAAGTTAAGTCATTCTGCATAATCAAGGTACTCAGCTTGCTGAA	149
Qy	549	GACATGACGCTGTATGTCTTTTTCAGCTGCAGACAGAGCTTTTTTAGCGAACATATG	608
Db	148	TGCTTATCTCTTCTAAGTCTAATTGAGCAATATGAATAATTTTGGATGGAGATCTCATC	89
Qy	609	TTAACTTTTTCATCTAGCTTTGCTGTTTGTGTACAAATGAAGAGCAGTCAAGAGGT	668
Db	88	TTAAATGTTTCCCTCTACCTACTTCTTAAATGTCCTATCAATGCAATGAACAATTT	29
Qy	669	GAATGA	674
Db	28	CCTTGA	23
RESULT 8			
US-10-793-626-4398			
; Sequence 4398, Application US/10793626			
; Publication No. US20050255478A1			
; GENERAL INFORMATION:			
; APPLICANT: KIMMERLY, WILLIAM JOHN			
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS			
; FILE REFERENCE: PU3480US			
; CURRENT APPLICATION NUMBER: US/10/793,626			
; PRIOR FILING DATE: 2004-03-04			
; PRIOR APPLICATION NUMBER: 60/164,258			
; PRIOR FILING DATE: 1999-11-09			
; NUMBER OF SEQ ID NOS: 4472			
; SOFTWARE: PatentIn Ver. 2.1			
; LENGTH: 2436			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: synthetic			
; OTHER INFORMATION: nucleic acid sequence			
US-10-793-626-4398			
Query Match			
Best Local Similarity 3.6%; Score 37.6; DB 8; Length 2436;			
Matches 55; Conservative 0; Mismatches 29; Indels 0; Gaps 0;			
Qy	972	ATCGAAGATGAGGACGAGCATTTCTCGTCAAAATTTCCCGGCTAGGCAAAAAACGGCA	1031
Db	2	ATCGAANAATGAATGATGCTTATTTAACAACAATTTCTCTGGATAGGAAAGAAACTGCA	61
Qy	1032	AGGCAGATCATCTTGCACCTGAAA	1055
Db	62	AGACAAATGTGTAGATTAAAA	85
RESULT 9			
US-09-925-065A-295479/c			
; Sequence 295479, Application US/09925065A			
; Publication No. US20040181048A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, David G.			
; TITLE OF INVENTION: Identification and Mapping of Single			
; FILE REFERENCE: 10827.135			
; CURRENT APPLICATION NUMBER: US/09/925,065A			
; PRIOR FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/243,096			
; PRIOR FILING DATE: 2000-10-24			
; PRIOR APPLICATION NUMBER: US 60/252,147			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: US 60/250,092			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: US 60/261,766			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/289,846			
; PRIOR FILING DATE: 2001-05-09			
; NUMBER OF SEQ ID NOS: 957086			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 295479			
; LENGTH: 644			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-065A-295479			
Query Match			
Best Local Similarity 3.5%; Score 36.8; DB 6; Length 644;			
Matches 101; Conservative 1; Mismatches 108; Indels 0; Gaps 0;			
Qy	527	AACATCATGTAGAAGATATGAACACATGACGCTGTTATGCTTTTTCAGCTGCAGACAG	586
Db	519	AAGAGCTTTTCGAGAAATGTGAACAATGCCACATTTTCTAATTTTGTGTTTGGAAAA	460
Qy	587	AAGCTTTTTTAGCGAACATATGTTAACTTTTTTTCATCTAGCTTTGCTGTTGTAC	646
Db	459	AACTTTTGTGTAATAATATGTTATTATTAACATGTAATAATGTTTATGATTGTTATTTT	400
Qy	647	AATGAAGAGCAGTCAAGAGGTGAATGACGTTGATCGAATTCGTAAGGAGGACATGGA	706
Db	399	AAATAAATTAATAAACAATGTTTGAATTTGTTTAAATTTCTAATATGTTAAATTTT	340
Qy	707	TTATGTATCGCCCAATATATTTGTCATTGA	736
Db	339	TTATTATKATCTTTTATTTTGTGAGACTGA	310



RESULT 10

US-11-117-187-209/c

Sequence 209, Application US/11117187

Publication No. US2005026560A1

GENERAL INFORMATION:

APPLICANT: PREUSS, DAPHNE

APPLICANT: COPENHAVER, GREGORY

TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS

FILE REFERENCE: ARCD:309US

CURRENT APPLICATION NUMBER: US/11/117,187

CURRENT FILING DATE: 2005-04-28

PRIOR APPLICATION NUMBER: US/09/531,120

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/125,219

PRIOR FILING DATE: 1999-03-18

NUMBER OF SEQ ID NOS: 212

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 209

LENGTH: 611587

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-11-117-187-209

Query Match 3.5%; Score 36.8; DB 12; Length 611587;

Best Local Similarity 51.9%; Pred. No. 96; Mismatches 0; Gaps 0;

Matches 83; Conservative 0; Indels 77; Indels 0; Gaps 0;

Qy 139 CAAATCCAGCTTGAAGAAGGTTTATCTCGACGGAGACGTTTGGAAATTGAGAATAAAGTAGAG 198

Db 89346 CAAATGATGTTAAATAATGCGTTTCTTCAAGGAGAACTCGAAGAAAAGCTGTACATGAAG 89287

Qy 199 GCGCTCGCACACTGAAGACTTTAAAGTGTCTTATAAGGGTGGCAGCTCATCGATCAG 258

Db 89286 CCACCAACCGGCTAGAGACATTAATGTCCCAATAAAGTCTTTAAGCTTTAAGAAAGCC 89227

Qy 259 AAAAAGGGGTTTATTCTGTTTCGAAACAGGTGGACGACA 298

Db 89226 ATCTACGGTTTAAGCAATCTCCGAGGCGTGTACCACA 89187

RESULT 11

US-09-925-065A-174583/c

Sequence 174583, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 174583

LENGTH: 364

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-174583

Query Match 3.5%; Score 36.6; DB 6; Length 364;

Best Local Similarity 55.0%; Pred. No. 5.4; Indels 2; Gaps 1;

Matches 93; Conservative 0; Mismatches 74; Indels 2; Gaps 1;



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:32:16 ; Search time 233 Seconds  
(without alignments)  
8048.616 Million cell updates/sec

Title: US-10-510-408-1  
Perfect score: 1055  
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatcttgacctgaaa 1055

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCUTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RS COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103.4	9.8	615	3	US-09-134-001C-780
2	89.8	8.5	14273	3	US-08-961-527-40
3	89	8.4	7159	3	US-08-956-171E-302
4	89	8.4	7159	3	US-08-781-986A-302
5	88.2	8.4	609	3	US-09-107-433-1114
6	87.4	8.3	594	3	US-09-583-110-2142
7	74.4	7.1	696	3	US-09-134-000C-3189
8	65.6	6.2	633	3	US-09-489-039A-6958
9	61.4	5.8	624	3	US-09-540-236-555
10	61.4	5.8	45613	3	US-09-596-002-22
11	60.2	5.7	609	3	US-09-328-352-2228
12	60.2	5.7	657	3	US-09-543-681A-361
C 13	50.6	4.8	3250	3	US-09-221-017B-960
C 14	49	4.6	1830121	3	US-09-557-884-1
C 15	49	4.6	1830121	3	US-09-643-950A-1
C 16	49	4.6	1830121	3	US-10-158-865-1
C 17	47.8	4.5	510	3	US-09-252-991A-13007
C 18	47.8	4.5	873	3	US-09-252-991A-12464
C 19	47.8	4.5	7218	2	US-08-232-463-14
C 20	46	4.4	609	3	US-09-902-540-2478
C 21	46	4.4	15209	3	US-09-902-540-1110
C 22	43.4	4.1	558	3	US-09-583-110-2141
C 23	40.6	3.8	79350	3	US-09-949-016-12467
C 24	40.6	3.8	79351	3	US-09-949-016-16275

ALIGNMENTS

RESULT 1

US-09-134-001C-780  
; Sequence 780, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 780  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-780

Query Match	9.8%	Score 103.4;	DB 3;	Length 615;
Best Local Similarity	56.4%	Pred. No. 3.8e-23;		
Matches	216;	Conservative	0;	Mismatches 161;
				Indels 6;
				Gaps 1;
QY	679	TCATCGAATTCGTAAGGAGGAGGATGATTATGATATCGCCCAATATATTGTCATTGAAA	738	
DB	14	TGTATGCATATATTAAAGGAAACATTATCTCAGTTGTTCCCTACACATGATAGTGGTTGAGA	73	
QY	739	ACGGCGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTAAGAAACACGCA	798	
DB	74	CATGCGCATAGTTATGAGATACAAACGCTTAATCTCTGTTTCAAAAATATCTTG	133	
QY	799	AAGAAACATC-----TATACATACCATTTATGTAAGAGAGACAGAAATGCGGTGACG	852	
DB	134	AAAAAGAGTCCAAATTTATCTTCACTTAATTTGTACGAGAGATGCTCAACTACTATATG	193	
QY	853	GCTTTTCACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGATCG	912	
DB	194	GCTTTTAAATCAAGAGGAGAAAAGAAATGTTCTTCTAGCTTAATAAAGAGTACTGGGATAG	253	
QY	913	GCCAAAAGGAGCGCTTCGCGATCCCTCGCTCCGGCGATCCGGGCGGTGATTGAAGCGA	972	
DB	254	GACCTAAATCAGCTTTAGCGATCTTCTTCAAGTACACCATCAATGAAGTTAAATTTGGCTA	313	
QY	973	TCGAGATGAGGACGAAAGCATTTTCTCGTCAAAATTTCCCGGGCGTAGGCAAAAACGGCAA	1032	

Db 314 TCGAAATGAAATGATCTTATTTAAACAAATTTCTGGGATAGGAAGAAACTGCAA 373  
QY 1033 GCGAGATCATCTTGACCTGAAA 1055  
Db 374 GACAAATTGTGTAGATTAAAA 396

RESULT 2  
US-08-961-527-40  
; Sequence 40, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14273 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-40

Query Match 8.5%; Score 89.8; DB 3; Length 14273;  
Best Local Similarity 54.0%; Pred. No. 7.6e-18;  
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY 675 ACGTTGATCGAATTCGTAAGGAGGAGATTGATTGATGATCGCCCAATATATTGTCATT 734  
Db 11348 ACTATGTAGCATATTTAAAGAGATCAATACCAAAATCTGCGCAATACATTTGCTT 11407

QY 735 GAAACGCGGGATCGGCTAT-----CAGATCTTCACGCCAAATCCGTTTATTATAAG 788  
Db 11408 GAAACCAATGGTATTGGTTATATCTGCAATGTGGCAATCTTATGCTTATTCAGGTCAG 11467

QY 789 AAAACAGCAAGAAACATCTATACATACCATTTATGTAAGAGAGACACGATCGGCTG 848  
Db 11468 GTTAATACGAGGCTCAGATTTATGTCATCAGGTTGTGCGTGAGGACGCCAATTTGCTT 11527

QY 849 TACGGCTTTTCACAGGGAAGAAAAATGCTGTTTACGAAAATGCTGAATGCTTACGGGG 908  
Db 11528 TATGATTTTCGTCAGAGGATGAGAAAAGCTCTTTCTAGTCTGATTCGCTCTCTGGG 11587

QY 909 ATCGGCCCCAAAAGGACGCTTCGATCTCTCGCTTCGGCGATCCGGGACGGGTGATGAA 968  
Db 11588 ATTGGTCTGTATCAGCTCTTCTGCTATTATCGCTGCTGATGACAAATGCTGGCTTGGTTCAA 11647

QY 969 GCGATCGAGATGAGGACGAGCATTTCTCGTCAAAATTTCCCGCGTAGGCAAAAAACG 1028  
Db 11648 GCCATTGAACCAAGAAATCATCCTACTTGACCAAGTTCCCTAAATTTGGCAAGAAACA 11707

QY 1029 GCAAGGCGAGATCATCTTGTGACCTGAAA 1055  
Db 11708 GCCCAGCAGATGGTGTGCTGGACTTGAA 11734

RESULT 3  
US-08-956-171E-302  
; Sequence 302, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,171E  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hyman  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224  
; TELEFAX: (301) 309-8439  
; INFORMATION FOR SEQ ID NO: 302:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7159 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 302:  
US-08-956-171E-302

Query Match 8.4%; Score 89; DB 3; Length 7159;  
Best Local Similarity 54.0%; Pred. No. 9.1e-18;  
Matches 207; Conservative 0; Mismatches 170; Indels 6; Gaps 1;

QY 679 TGATCGAATTCGTAAGGAGCGATTGATTGATATCGCCCAATATATTGTCATTGAAA 738  
Db 3236 TGTACGCGTGTGTCAAAGGTAAGTTAACACATTTATATATCTTACACACGCTAGTTGTTGAAA 3285

QY 739 ACGCGGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTATTATAAGAA-----AA 792  
Db 3286 CTGCTGGTGTGGTTATGAAATTCACACCAAAATCTTATCGTTTTCAAAAGCATCTAG 3345

QY 793 ACAGCAAGAAACAAATCTATACATACCATTTATGTAAGAGAGACACGAATGCGCTGACG 852

Db 3346 ATCATGAAGTTTAAATTCATACATCTTTAAATGTTTCGTAAGATGCAACAATTTGTATG 3405  
Qy 853 GCTTTTCGACAAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCG 912  
Db 3406 GATTTAGTAGTGAAGAGAGAAAGATAATGTTCTTGAGTTTAATTAAGTTACTGATATG 3465  
Qy 913 GCCAAAAGGAGCGCTTGGGATCCCTGCTTCCGGGATCCGGGAGCGGTGATTTGAAGCGA 972  
Db 3466 GTCCGAAATCAGCTTTAGCTATTTTAGCGCAAGTACGCTTAATGAAGTAAACGTGCGCA 3525  
Qy 973 TCGAAGTAAGGACGACGATTTCTCGTCAAAATTTCCGGGCTAGGCAAAAAACGGCAA 1032  
Db 3526 TTGAATAATGAATAATGATAGTATTTAACTAAATTTCCAGGAATTTGTAAGAAAAACGGCAA 3585  
Qy 1033 GCGAGATCATCTTGACCTGAAA 1055  
Db 3586 GACAGATTGCTCTTAGATTTAAAA 3608

## RESULT 4

US-08-781-986A-302  
; Sequence 302, Application US/08781986A  
; Patent No. 6737248

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 302:

SEQUENCE CHARACTERISTICS:

LENGTH: 7159 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-302

Query Match 8.4%; Score 89; DB 3; Length 7159;  
Best Local Similarity 54.0%; Pred. No. 9,1e-18;  
Matches 207; Conservative 0; Mismatches 170; Indels 6; Gaps 1;

Qy 679 TGATCGAATTCGTAAGGAGCGATTTGATATGATCGCCCAATATATTTGTCATTTGAAA 738

Db 3226 TGTACCGGTATGCAAGGTAAGTTAAACACATTTATATCTACACACGTAGTTGTTGAAA 3285

Qy 739 ACGGCGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTATTTATAGAA-----AA 792

Db 3286 CTGCTGGGTGGTTATGAAATTTCAACACCAAAATTTCTTATCGTTTTCAAAAGCATCTAG 3345

Qy 793 ACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGAAAGACACGAATGCGCTGTACG 852  
Db 3346 ATCATGAAGTTTAAATTCATACATCTTTAAATGTTTCGTAAGATGCAACAATTTATTGTATG 3405  
Qy 853 GCTTTTCGACAAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCG 912  
Db 3406 GATTTAGTAGTGAAGAGAGAAAGATAATGTTCTTGAGTTTAATTAAGTTACTGATATG 3465  
Qy 913 GCCAAAAGGAGCGCTTCCGATCTCGCTTCCGGCGATCCGGGAGCGGTGATTTGAAGCGA 972  
Db 3466 GTCCGAAATCAGCTTTAGCTATTTTAGCGCAAGTACGCTTAATGAAGTAAACGTGCGCA 3525  
Qy 973 TCGAAGTAAGGACGACGATTTCTCGTCAAAATTTCCGGGCTAGGCAAAAAACGGCAA 1032  
Db 3526 TTGAATAATGAATAATGATAGTATTTAACTAAATTTCCAGGAATTTGTAAGAAAAACGGCAA 3585  
Qy 1033 GCGAGATCATCTTGACCTGAAA 1055  
Db 3586 GACAGATTGCTCTTAGATTTAAAA 3608

## RESULT 5

US-09-107-433-1114

; Sequence 1114, Application US/09107433

; Patent No. 6800744

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

THERAPEUTICS

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1114:

SEQUENCE CHARACTERISTICS:

LENGTH: 609 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...609

FOR DIAGNO

```
SEQUENCE DESCRIPTION: SEQ ID NO: 1114:
US-09-107-433-1114

Query Match      8.4%; Score 88.2; DB 3; Length 609;
Best Local Similarity 53.7%; Pred. No. 3.6e-18;
Matches 208; Conservative 0; Mismatches 173; Indels 6; Gaps 1;

Qy 675 ACGTGTGATCGAATTCGTAAGAGGACGATTGATTATGATCGCCCAATATATTGTCTATT 734
Db 13 ACTATGTACGCATATTTAAAGGAATCATTAACCAAAATTTACTGCAATACATTTGTTCTT 72

Qy 735 GAAAAACGGCGGATCGGCTAT-----CAGATCTTACGCCCAATCCGTTTATTATAAG 788
Db 73 GAACCAATGATTTGGTTATATCTCGATGTGGCAATCTTATGCTATTTCAGGTCAG 132

Qy 789 AAAAACAGCAAGAAACAATCTATACATACCATATGTAAAGAGACAGCAATGCGGTG 848
Db 133 GTTAATCAGGAGGCTCAGATTTATGTGCATCAGGTTGTGCGTGAGGACGCCCAATTTGCTT 192

Qy 849 TACGGCTTTTCGACNAGGAGAAATAATGCTGTTTACGAAATGCTGAATGTTACGGG 908
Db 193 TATGATTTTCGCTCAGAGGATGAGAAAGGCTCTTTCTTATGCTAAATTTTCGGTCTCTGGG 252

Qy 909 ATCGGCCCAAGAGCGCTTTCGATCCTCGCTTCCGGCGATCCGGGACGCGTGAATTGAA 968
Db 253 ATTGGCTCTGTATCAGCTCTTGCTATTATCTGCTGATGACATGCTGGCTTGTTCAA 312

Qy 969 GCGATCGAGATGAGGACGAAGCAATTTCTCGTCAAAATTTCCCGGCTAGGCAAAAAACG 1028
Db 313 GCCATTGAACCAAGAACATCACCTACTTACCAAGTTCCCTAAATTTGGCAAGAAACA 372

Qy 1029 GCAAGGCGATCATCTTTCGACCTGAAA 1055
Db 373 GCCCAGCAGATGGTCTGACTTGGAA 399

RESULT 6
US-09-583-110-2142
; Sequence 2142, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2142
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2142

Query Match      8.3%; Score 87.4; DB 3; Length 594;
Best Local Similarity 53.8%; Pred. No. 6.5e-18;
Matches 206; Conservative 0; Mismatches 171; Indels 6; Gaps 1;

Qy 679 TGATCGAATTCGTAAGAGGACGATTGATTATGATCGCCCAATATATTGTCAATTGAAA 738
Db 2 TGTACGCATATTTAAAGGAATCATTAACCAAAATTTACTGCCAAATACATTTGTTCTTAAA 61

Qy 739 ACGGCGGATCGGCTAT-----CAGATCTTACGCCCAATCCGTTTATTATAAGAAA 792
Db 62 CCAATGGTATTTGTTTATATCTCGCATGTGGCCAATCTTATGCTCTATTTCAGGTCAGGTTA 121

Qy 793 ACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGACACAGCAATGCGCTGTACG 852
```

```

122 ATCAGGAGGCTCAGATTTATGTGCATCAGGTTGTGCGTAGGAGCGCCCAATTTGCTTTATG 181
Qy 853 GCTTTTCGACAAGGAGAAATAATGCTGTTTACGAAAAATGCTGAAATGTTACGGGATCG 912
Db 182 GATTTTCGCTCAGAGGATGAGAAAAAGCTCTTCTTACTGCTAAATTTTCGGTCTCTCGGATG 241
Qy 913 GCCCAAAAGAGCGCTTTCGATCCTCGCTTCCGCGATCCGGAGCGGTGATTCAAGCGA 972
Db 242 GTCTGTATCAGCTCTTGTCTATTATCGCTGCTGATGACAATGCTGGCTTGGTTCAAGCCA 301
Qy 973 TCAGAAATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGCGTAGGCAAAAAACGGCAA 1032
Db 302 TTGAAACCAAGACATCACCTACTTGACCAAGTTCCCTAAATTTGCAAGAAACAGGCC 361
Qy 1033 GGCAGATCATCTTTCGACCTGAAA 1055
Db 362 AGCAGATGGTCTGCACTTGGAA 384

RESULT 7
US-09-134-000C-3189
; Sequence 3189, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3189
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3189

Query Match      7.1%; Score 74.4; DB 3; Length 696;
Best Local Similarity 53.4%; Pred. No. 1.3e-13;
Matches 206; Conservative 0; Mismatches 171; Indels 9; Gaps 2;

Qy 679 TGATCGAATTCGTAAGAGGACGATTGATTATGATCGCCCAATATATTGTCAATTGAAA 738
Db 83 TGTACGAATATATTTATTGGAAGTAACGTTTGTAGCCCATACTATATTGCTGTTGAAA 142
Qy 739 ACGGCGGATCGGCTATCAGATCTTTCACGCCAAATCGTTTATTATTTAT-----AAGAAA 792
Db 143 CAAATGGCAITGGCTATCAGATTTTCAGTCGATAACCGTATCGCTATTTCAGAAAAAATGG 202
Qy 793 ACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGACACAGCAATGCGCTGTACG 852
Db 203 ATACGGATTTAAACTATATCTTCCCAAGTCGTGCGGAAGATGCACACTTTTATTG 262
Qy 853 GCTTTTCGACAAGGAGAAATAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCG 912
Db 263 GTTTTGGCAGTTTAGAAGAAAAACAATTTGTTCTTAAAAATTAATGATGCTCTCGGAATG 322
Qy 913 GCCCAAAAGGACGCTTTCGATCCTCGCTTC-----CGCGATCCGGAGCGGTGATTGAAG 969
Db 323 GACCTAAAGTGGCTTAGCTATCATGCTTCTGTGGCTGATCATGTTGGTTTAAATCAATG 382
Qy 970 CGATCGAATGAGGACGAAGCAATTTCTCGTCAAAATTTCCCGGCTAGGCAAAAAACGG 1029
Db 383 CAATTGAAGGGAAGATGTTACATATTTAACGAAGTTCCCAAGGTTTCCGCAAGAAACAG 442
Qy 1030 CAAGGCGATCATCTTTCGACCTGAAA 1055
Db 443 CCCAACAAATGATTTTAGATTTAAAA 468
```

RESULT 8

US-09-489-039A-6958

; Sequence 6958, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 6958

; LENGTH: 633

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-6958

Query Match 6.2%; Score 65.6; DB 3; Length 633;

Best Local Similarity 54.0%; Pred. No. 9.6e-11;

Matches 134; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 808 TCTATACATACCAATTATGTAAGAGAAGACACGATGCGCTGACGGCTTTTCGACAAGG 867

DB 161 TCTTACCACCTTTGTGTGCGTGAAGATGCCAGCTGCTGTATGGGTTTTAACACAACAAC 220

QY 868 AAGAAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCGC 927

DB 221 AGGAACGTACCTCTTTAAGAGCTGATTAACCAACAGCGGTGGGCCCCAGCTGGCGC 280

QY 928 TTGGGATCTCGCTTCCGGCGATCCGGAGCGGTGATTGAAGCGATCGAAGATGAGGACG 987

DB 281 TGGCGATCTCTCCGGCATGTTCGGCGCAGCAGTTCTGTCACCGCGTTGAACGGGAAGG 340

QY 988 AAGCATTTCTGTCAAATTTCCGGCGCTAGCGAAAAAGCGAAGCAGATCATCTTTCG 1047

DB 341 TCGCTCGTGTGAGCTGCGGGGATTGCGAAAAAACCGCTGAGCGTCTGATCGTTG 400

QY 1048 ACCTGAAA 1055

DB 401 AGATGAAA 408

RESULT 9

US-09-540-236-555

; Sequence 555, Application US/09540236

; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 555

; LENGTH: 624

; TYPE: DNA

; ORGANISM: M.catarrhalis

US-09-540-236-555

Query Match 5.8%; Score 61.4; DB 3; Length 624;

Best Local Similarity 53.0%; Pred. No. 2.3e-09;

Matches 131; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 808 TCTATACATACCAATTATGTAAGAGAAGACACGATGCGCTGACGGCTTTTCGACAAGG 867

DB 149 TCTATGACATCATCATGTCGGAAGATGCCATAAATTGTTGGTTTTTATGATCGTC 208

QY 968 AAGAAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCGC 927



; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GT99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2228  
; LENGTH: 609  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-228

Query Match 5.7%; Score 60.2; DB 3; Length 609;  
Best Local Similarity 51.7%; Pred. No. 5.8e-09;  
Matches 137; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 787 AGAAACAGCAAGCAAAATCTATACATACCATATGTAAAGAGACACGAAATGGCG 846  
Db 125 AAAAAGGCCAAGATTACGCTATGACGCGATTTGGTGGTGAAGATGCTCAGCAGC 184  
Qy 847 TGTACGGCTTTTCGACAGGGAAGAAAAATGCTGTTTACGAAAATGCTGAATGTTACGG 906  
Db 185 TTTATGGTTTATGATGCTCTCAAGAGAAAACCTATTTCCGACACCTTACTAAAAGTAAATG 244  
Qy 907 GGATCGGCCCAAGAGCGCTTTCGATCCTCGCTTCGGCGATCCGGAGCGGTGATTG 966  
Db 245 GTGGGGCCCAAAATGCGCATAGTATCTTTCTACGTTAAGTGTGAATGCTGGTAC 304  
Qy 967 AAGCGATCGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAAA 1026  
Db 305 ACACAAATGAACATGACGATGTAATACTTTGGTTAAAGTTCCGGGTGTAGGCAAAAAA 364  
Qy 1027 CGGCAAGCGATGATCCTTGGACT 1051  
Db 365 CAGCTGAACGTTTAAATGATTGAAC 389

## RESULT 12

US-09-543-681A-361  
; Sequence 361, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 361  
; LENGTH: 657  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-361

Query Match 5.7%; Score 60.2; DB 3; Length 657;  
Best Local Similarity 52.6%; Pred. No. 5.8e-09;  
Matches 131; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 807 ATCTATACATACCATATGTAAGAGACACAGCAATGCGGTGTACGGCTTTTCGACAGG 866  
Db 175 ATCTACACACAAATTTATTGTCGTGAAGATGCGCAATTACTATATGTTTAAATCAAAAA 234  
Qy 867 GAAGAAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGATCGGCCCAAAAGGAGCG 926  
Db 235 CAAAGACGGGCATGTTTCGTGAATGTAATAAGTCAATGTTGTTAGGCGCTAAGCTAGCA 294  
Qy 927 CTTGCGATCCTGCTTCGGCGCATCCGGAGCGGTGATTGAAGCGATCGAAGATGAGGAC 986  
Db 295 CTGGCTATTTATCAGGAATGCTCGACGCCCAATTTGTTCACCGCAATTTGAAATGAATCAATCC 354

Qy 987 GAAGCATTTTCGTCAAATTTCCGGCGTAGGCAAAAAAAGCGAAGCGACATCATCTT 1046  
Db 355 ATCTCTCATATTAGTTAAATTAACCGGAGTCGTTAAAAAACCAGACGATTAGTGTT 414  
Qy 1047 GACCTGAAA 1055  
Db 415 GAAATGAAA 423

## RESULT 13

US-09-221-017B-960/c  
; Sequence 960, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: ROSE, BRUCE C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSEQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/221,017B  
; FILING DATE: 23-DEC-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP1182  
; FILING DATE: 31-DEC-1997

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP1546  
; FILING DATE: 30-JAN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP2911  
; FILING DATE: 09-APR-1998

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU98/01023  
; FILING DATE: 10-DEC-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monroy, Gladys H  
; REGISTRATION NUMBER: 32,430  
; REFERENCE/DOCKET NUMBER: 27340-20021.00

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-813-5600  
; TELEFAX: 650-494-0792  
; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 960:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3250 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double

; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: UNKNOWN

; ORIGINAL SOURCE:  
; ORGANISM: PORYPHYROMONAS GINGIVALIS  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1...3250

US-09-221-017B-960

Query Match 4.8%; Score 50.6; DB 3; Length 3250;  
Best Local Similarity 49.1%; Pred. No. 2.2e-05;  
Matches 134; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 783 TATAGAAAACAGCAAGAAACAAATCTATACATACATATATGTAAGAGAACACAGCAAT 842  
|||||  
Db 1540 TATACGGGGAAGAAAGAGAGGACTTATTTGGATTACACACTGATCCGAGAGATGCCCAT 1481  
|||||  
QY 843 GCGCTGTACGGCTTTTCGACAAGGAAGAAATGCTGTTTACGAAATGCTGAATGTT 902  
|||||  
Db 1480 TTATTGTATGGCTTTTCACGAAGAGAGCGTACGCTCTTCGGCAACTCACATCTGTC 1421  
|||||  
QY 903 ACGGGATCGGCCCAAAAGAGCGCTTGCATCTCGCTTCGGCGGATCCGGAGCGGTG 962  
|||||  
Db 1420 AGCGGTGTCGGGCTACGACGCGCACAGCTCATCTTCTCTATGCTCTCTCAAGAGCTG 1361  
|||||  
QY 963 ATTGAAGCATCGAGATGAGACGAGCAATTTCTCGTCAAAATTTCCCGGCGTAGGCAAA 1022  
|||||  
Db 1360 GCCGCACTATTACACAGGCGGCGGCTGAAAGCGGTGAAGGCGCATCGGTCTG 1301  
|||||  
QY 1023 AAAACGGCAAGGCAGATCATCTTGACCTGAAA 1055  
|||||  
Db 1300 AAGACCGCTCAGCGTATCATCTGATCTGAAA 1268  
|||||

## RESULT 14

US-09-557-884-1/c  
; Sequence 1, Application US/09557884  
; Patent No. 6506581  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1  
Query Match 4.6%; Score 49; DB 3; Length 1830121;  
Best Local Similarity 49.4%; Pred. No. 0.0036;  
Matches 127; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 799 AAGAACAACTCTATACATACATATGTAAGAGAGACAGCAATCGCTGACGCTTTT 858  
|||||  
Db 344955 AAACAACTTTATTTCACCCCATCTTGTGTTCGGGAAGATGCTCATTACTTCTTTGGATTG 344896  
|||||

QY 859 CGACAAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCAA 918  
|||||  
Db 344895 CCAGAAACAGACCGCACTTTATTTCGTGAATTAATTAACAAATGCGGTGGGCGCTA 344836  
|||||  
QY 919 AAGGAGCGCTTGGATCTCGCTTCGGGGATCCGGAGCGGTGATTGAAGCGATCGAGA 978  
|||||  
Db 344835 AATTAGCCTTAGCGATTTTATCCGCCATGTCGGTCGAACAAATTTGCCTATGCAATAGAGA 344776  
|||||  
QY 979 ATGAGAGCAAGCAATTTCTCGTCAAAATTTCCCGGCGTAGGCAAAAAACGCAAGGACAGA 1038  
|||||  
Db 344775 GAGAGAACTTCTAAACTTACTAAATTCAGGTTTGGCAAAAAACAGCTGAACGTT 344716  
|||||  
QY 1039 TCATCCTTGACCTGAAA 1055  
|||||  
Db 344715 TGTAGTTGAGCTCAAA 344699  
|||||

## RESULT 15

US-09-643-990A-1/c  
; Sequence 1, Application US/09643990A  
; Patent No. 6528289  
; GENERAL INFORMATION:  
; APPLICANT: Robert D. Fleischmann  
; Mark D. Adams  
; Owen White  
; Hamilton O. Smith  
; J. Craig Venter  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB186PIC1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1  
Query Match 4.6%; Score 49; DB 3; Length 1830121;  
Best Local Similarity 49.4%; Pred. No. 0.0036;  
Matches 127; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy	799	AAGAAACAATCTATACATACCATTATGTAGAGAGAGACGAAATGCGCTGTACGGCTTTT	858
Db	344955	AAACAACTTTATTCAACCATCTTGTGTTTCGGGAAGATGCTCACTTACTCTTTGGATTGG	344896
Qy	859	CGACAAGGGAGAGAAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCAA	918
Db	344895	CCGAGAAAACAGACCGCACCTTATTTTCGTGAATTAAATTAACAAATGGGGTGGGGCCTA	344836
Qy	919	AAGGAGCGCTTCCGATCCTCGCTTCCGGGATCCGGGAGCGGTGATTGAAGGATCGAGA	978
Db	344835	AATTAGCCTTAGCGATTTTATCCGCCATGTCCGGTCGAAACAATTTGCCCTATGCAATAGAGA	344776
Qy	979	ATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGGGTAGGCAAAAAACGGCAAGGCAGA	1038
Db	344775	GAGAAGAACTTTCTAAACTTACTAAAAATCCAGGTGTTGGCAAAAAACAGCTGACGTT	344716
Qy	1039	TCATCCTTGACCTGAAA	1055
Db	344715	TGTTAGTTGAGCTCAAA	344699

Search completed: February 27, 2006, 06:36:34  
 Job time : 245 secs